EDITOR: Joshua J. Yi, josh@joshuayipatentlaw.com

DEPARTMENT: MICRO LAW

Does Academic Research Drive Industrial Innovation in Computer Architecture?—Analyzing Citations to Academic Papers in Patents

Joshua J. Yi, *The Law Office of Joshua J. Yi, PLLC, Austin, TX, 78750, USA*

In my prior article,[1] I analyzed the 1) number of "backward" citations to previously issued/published U.S. patents and publications, foreign patents, and other publications, and 2) the number of "forward" citations, by later-filed patents, for patents that were issued to 18 leading computer architecture companies which were filed between 1996 and 2020. In this article, I extend that prior work by further analyzing a subset of other publications, namely, academic papers. Analyzing the number of citations by issued patents to papers published in the top computer architecture conferences may reveal whether academic research is driving industrial innovation in computer architecture. More specifically, in this article, I examine the number of citations to papers published in International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), International Symposium on High-Performance Computer Architecture (HPCA), International Symposium on Computer Architecture (ISCA), and International Symposium on Microarchitecture (MICRO). Given that these four conferences are the most prestigious in computer architecture, they are an excellent proxy for academic research as they likely contain the most innovative ideas. As such, papers from these conferences could provide important building blocks for a company's patented inventions; if so, these papers should be cited in the company's patents.

CITATIONS TO TOP COMPUTER ARCHITECTURE CONFERENCES

On the first page of each U.S patent is a section titled "References Cited." The applicant or the examiner provides the references in the references cited section. The applicant has a duty to disclose to the U.S. Patent and Trademark Office ("PTO") all prior art and other information "material to patentability."[a] An applicant's failure to fulfill this duty could result in the patent being unenforceable. In addition to this duty, by providing more of the relevant prior art to the examiner, the examiner is able to make a more informed decision as to whether the patent is different than the prior art, which may increase the likelihood that the ensuing patent is valid over prior art.

ANALYZING THE NUMBER OF CITATIONS IN ISSUED PATENTS TO PAPERS PUBLISHED IN THE TOP COMPUTER ARCHITECTURE CONFERENCES MAY REVEAL WHETHER ACADEMIC RESEARCH IS DRIVING INDUSTRIAL INNOVATION IN COMPUTER ARCHITECTURE.

The references cited section divides references into three groups: 1) U.S. patent documents, which includes both U.S. patents and U.S. patent publications, 2) foreign patents documents, and 3) "Other publications." U.S. patent publications are applications that are published 18 months after the earliest priority date of the patent. The other publications section is a catch-all category that may include dictionaries, manuals, standards, legal briefs, and, most relevant to this article, academic papers.

[a]37 C.F.R. § 1.56

0272-1732 © 2023 IEEE
Digital Object Identifier 10.1109/MM.2022.3226101
Date of current version 13 January 2023.

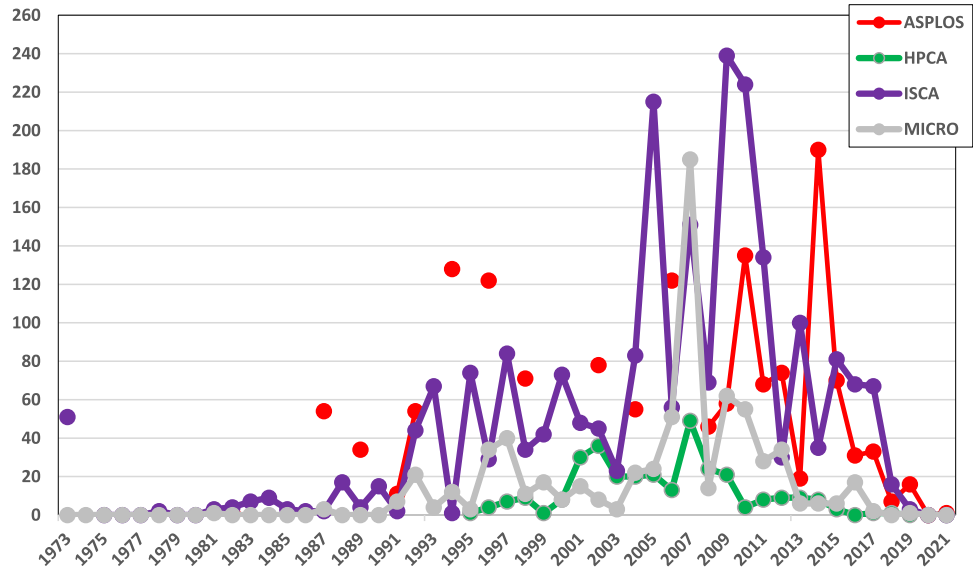

FIGURE 1. Citations in patents issued between December 28, 1977 and August 16, 2022 to papers published in ASPLOS, HPCA, ISCA, and MICRO.

To answer the question of whether academic research drives industrial innovation, I expanded my dataset from the approximately 500,000 patents that were issued to the 18 computer architecture companies to approximately 7.1 million U.S. patents that were issued between 28 December 1977 (U.S. Patent No. 4,000,000) and 16 August 2022 (U.S. Patent No. 11,419,251). For each conference (i.e., ASPLOS, HPCA, ISCA, and MICRO), I searched for both the full conference name and the acronym within the other references section. For each patent that cited a paper from ASPLOS, HPCA, ISCA, and MICRO, I also extracted the original assignee for that patent.

The results show that, out of 7.1 million issued patents, there were only 5,890 citations to 1,018 papers published in one of those four conferences. These results tend to indicate that academic research in computer architecture does not drive industrial innovation.[b,2]

While 5,890 patents represent 0.08% only of all 7.1 million issued patents, that number is misleading because it compares apples and oranges. More specifically, the number of computer architecture patents issued between 28 December 1977 and 16 August 2022 is probably between 300,000 and 400,000, which means that 1.47%–1.96% of issued computer architecture patents cite to a top computer architecture conference.[c] This seems like a very low percentage and may indicate at least one of two things. First, this may indicate that applicants are unable to find or are unaware of these papers. Second, more importantly, applicants may believe that papers published in top computer architecture conferences are not relevant to their patent applications. One reason why applicants may believe that is because they consider academic papers to be too impractical or too "blue-sky" to be relevant. Another reason may be that applicants may not necessarily appreciate the relevance given that academic papers typically focus on results and are page limited (thus resulting in limited disclosures of the proposed solution), as well as being directed toward Ph.D. students and above.

Figure 1 depicts the number of citations in issued patents between 28 December 1977 and 16 August 2022 to papers published in the proceedings of ASPLOS, HPCA, ISCA, and MICRO.

It is important to note that some conferences were not held every year. For example, ASPLOS did not start until 1982, and was not held in 1983 to 1986, 1993, 1995,

[b]Other studies have also shown "much of the contribution of knowledge spillovers from academia may be largely confined to bioscience-related inventions," and that the increase in information technology-related patents has not been driven by academic research.

[c]I intentionally tried to be conservative with my estimate of the number of issued computer architecture patents. To the extent it is actually higher, e.g., 600,000–800,000, the percentage of issued computer architecture patents that cite to a top computer architecture conference would be even lower, e.g., 0.74%–0.98%.

1997, 1999, 2001, 2003, 2005, and 2007. HPCA did not start until 1995. ISCA was not held in 1974.

The results in Figure 1 show several interesting results. First, ISCA had many more citations in the earliest years of that conference, as compared to MICRO, for the corresponding timeframes. More specifically, even though ISCA started after MICRO, ISCA had 130 citations to conferences held between 1973 and 1990 while MICRO only had four citations to proceedings during that timeframe. One paper in the first ISCA (ISCA 1973) was cited 51 times. The title of this paper was "DAP—A Distributed Array Processor" and was authored by Stewart F. Reddaway.

Second, the most highly cited paper was published in ASPLOS 2000, and was cited 1,070 times. The title of this paper was "OceanStore: An Architecture for Global-Scale Persistent Storage" and the authors of the paper were John Kubiatowicz, David Bindel, Yan Chen, Steven Czerwinski, Patrick Eaton, Dennis Geels, Ramakrishna Gummadi, Sean Rhea, Hakim Weatherspoon, Westley Weimer, Chris Wells, and Ben Zhao. To put this in context, the next closest paper was cited only 142 times. Furthermore, the OceanStore paper was cited more time twice as often as the total number of citations to HPCA and MICRO combined. Overall, the OceanStore paper accounts for approximately 18% of all citations to ASPLOS, HPCA, ISCA, and MICRO.

By comparison, according to ACM's digital library, as of 7 September 2022, the OceanStore paper was cited a similar number of times, 1,253. This made it the fourth most highly cited ASPLOS paper (the top three papers were cited 2,099, 1,409, and 1,351 times), again according to ACM's digital library.

One conclusion from these results is that inventors of issued patents and researchers publishing in academic journals and conferences have different views on what is relevant to their work. More specifically, the OceanStore paper was cited approximately eight times more than the next closest paper by issued patents. By contrast, this paper was cited approximately half as many times as the most highly cited ASPLOS paper in ACM's digital library. More generally, for papers cited at least 10 times in an issued patent (102 papers), a correlation analysis between 1) the number of times those papers were cited in an issued patent and 2) the number of citations for those papers according to ACM's digital library, yields a correlation value of 0.244. This relatively low, but positive, correlation value indicates inventors of issued patents and researchers publishing academic papers have a significant difference of opinion as to what is relevant to their work.

While the OceanStore paper was cited very frequently, it was cited by a relatively small number of assignees. More specifically, patents from IBM, Pure Storage, and Cleversafe account for 1,000 of the 1,070 citations; the remaining 70 citations were in patents to 12 other assignees. Similarly, the second most highly cited paper—"Leveraging 3D Technology for Improved Reliability," which was authored by Niti Madan and Rajeev Balasubramonian and which was published in MICRO 2007—was cited a total of 142 times. But all 142 citations were by a single assignee, Monolithic 3-D Inc. The third most highly cited paper—"eNVy: A Non-Volatile, Main Memory Storage System," which was authored by Michael Wu and Willy Zwaenepoel and which was published in ASPLOS 1994—was cited a total of 100 times. But of those 100 citations, patents from two companies—Intel and Google—accounted for 77 of those citations, while the remaining 23 citations were from a total of eight other companies.

These results indicate that a few assignees found the most highly cited papers to be very relevant while most assignees did not find these papers to be relevant. These results indicate that, at most, academic research drives industrial innovations only for a small set of companies.

ON THE OTHER HAND, A SHARP INCREASE IN THE CUMULATIVE PERCENTAGE OVER A SHORT PERIOD OF TIME COULD INDICATE AN INCREASING NUMBER OF PATENTS DIRECTED TOWARDS THE PAPER'S AREA OF TECHNOLOGY OR THAT THE IDEAS DISCLOSED IN THE PAPER MAY FINALLY BE RELEVANT TO COMPANIES.

Third, ASPLOS papers are generally more frequently cited than HPCA, ISCA, and MICRO papers. More specifically, ASPLOS papers were cited a total 2,616 times (44.3% of all citations), while HPCA, ISCA, and MICRO papers were cited 307 (5.2%), 700 (11.9%), and 2,267 times (38.4%), respectively. One reason why HPCA papers were not cited more often is because it is the "youngest" conference, with the first HPCA conference being held in 1995. On the other hand, ASPLOS was only held 27 times between 1973 and 2021, while ISCA and MICRO were held 48 and 49

TABLE 1. Top 10 most frequently cited ASPLOS, HPCA, ISCA , and MICRO papers in patents.

Citations	Conference	Title	Authors
1070	ASPLOS 2000	OceanStore: An Architecture for Global-Scale Persistent Storage	John Kubiatowicz, David Bindel, Yan Chen, Steven Czerwinski, Patrick Eaton, Dennis Geels, Ramakrishna Gummadi, Sean Rhea, Hakim Weatherspoon, Westley Weimer, Chris Wells, and Ben Zhao (all from the University of California, Berkeley)
142	MICRO 2007	Leveraging 3D Technology for Improved Reliability	Niti Madan and Rajeev Balasubramonian (all from the University of Utah)
100	ASPLOS 1994	eNVy: A Non-Volatile, Main Memory Storage System	Michael Wu and Willy Swaenepoel (all from Rice University)
98	ASPLOS 2014	DianNao: A Small-Footprint High-Throughput Accelerator for Ubiquitous Machine-Learning	Tianshi Chen, Zidong Du, Ninghui Sun, Jia Wang, Chengyong Wu, Yunji Chen (all from the Institute of Computing Technology, Chinese Academy of Sciences), and Olivier Temam (Inria)
80	ISCA 2011	Kilo-NOC: A Heterogeneous Network-on-Chip Architecture for Scalability and Service Guarantees	Boris Grot (University of Texas at Austin), Joel Hestness (University of Texas at Austin), Stephen W. Keckler (University of Texas at Austin, NVIDIA), and Onur Mutlu (Carnegie Mellon University)
80	ASPLOS 2010	Fairness via Source Throttling: A Configurable and High-Performance Fairness Substrate for Multi-Core Memory Systems	Eiman Ebrahimi (University of Texas at Austin), Chang Joo Lee (Intel), Onur Mutlu (Carnegie Mellon University), and Yale N. Patt (University of Texas at Austin)
79	ISCA 2010	Translation Caching: Skip, Don't Walk (the Page Table)	Thomas W. Barr, Alan L. Cox, and Scott Rixner (all from Rice University)
77	ISCA 2010	Aergia: Exploiting Packet Latency Slack in On-Chip Networks	Reetuparna Das (Pennsylvania State University), Onur Mutlu (Carnegie Mellon University), Thomas Moscibroda Microsoft Research), and Chita R. Das (Pennsylvania State University)
69	ISCA 2009	Scalable High Performance Main Memory System Using Phase-Change Memory Technology	Moinuddin K Qureshi, Vijayalakshmi Srinivasan, and Jude A. Rivers (all from IBM Research)
67	ISCA 2005	A High Throughput String Matching Architecture for Intrusion Detection and Prevention	Lin Tan and Timothy Sherwood (all from the University of California, Santa Barbara)

times, respectively. Therefore, the gap between ASPLOS and the other conferences could be even higher if ASPLOS was held every year between 1973 and 2021. Furthermore, the median number of citations for each conference, excluding the years where there were no citations, was 55 for ASPLOS, 9 for HPCA, 38.5 for ISCA, and 13 for MICRO. Similarly, the mean number of citations was 61.9 for ASPLOS (excluding the OceanStore paper), 13.3 for HPCA, 54.2 for ISCA, and 23.3 for MICRO.

These results may indicate that while applicants and applicant's attorneys may not frequently search through academic papers and/or find academic papers to be particularly relevant, when they do, they tend to cite to papers that were published in the most prestigious conferences, i.e., ASPLOS and ISCA. Therefore, to the extent that citations to ASPLOS, HPCA, ISCA, and MICRO are representative of citations to academic journals and conferences, these results might indicate that more prestigious conferences, i.e., ASPLOS and ISCA, tend to drive more industrial innovations, as compared to less prestigious conferences.

Table 1 lists the papers with the highest number of citations in issued patents.

Given that ASPLOS and ISCA papers account for over 82% of all citations, it is not that surprising to see that of the Top 10 most frequently cited papers, nine of them are from ASPLOS and ISCA.

Figure 2(a) and (b) depicts the percentage of the total number of citations for each Top 10 paper based on the filing year of the citing patents; note that the

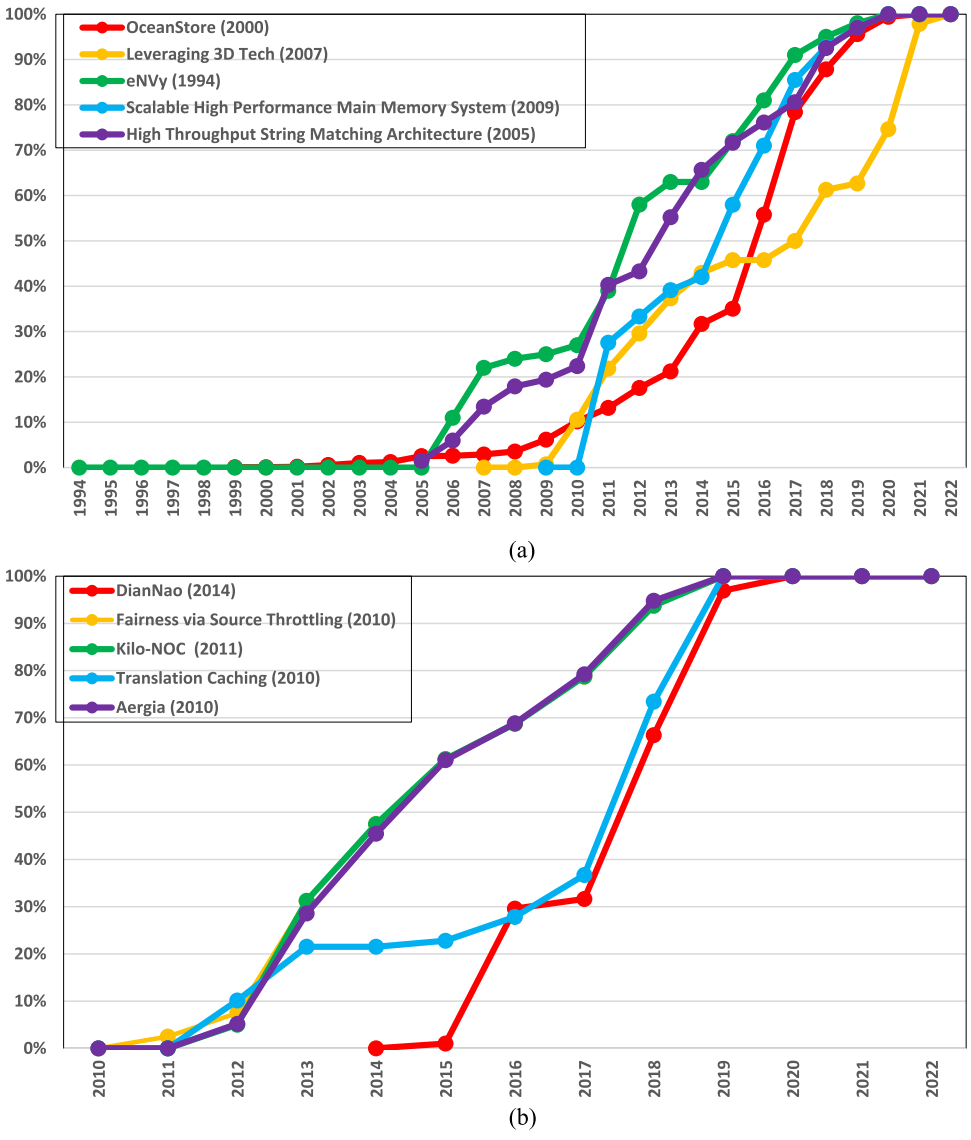

FIGURE 2. Percentage of the total number of citations each Top 10 paper was cited by patents filed in a particular year; percentages are cumulative. (a) Top 10 papers published 2009 and earlier. (b) Top 10 papers published 2010 and later.

percentage is a cumulative percentage in order to facilitate comparisons across papers. For example, the OceanStore paper was published in 2000. By 2010, it was cited by 109 patents that were filed in 2010 and earlier, which represents 10.2% of its 1,070 total citations. (Interestingly, although the OceanStore paper was published in 2000, the earliest filed patent that cited to the OceanStore paper was filed the year before it was published, in 1999. This indicates that the applicant of that patent found the OceanStore paper after the filing of that patent, but before it issued.) Figure 2(a) depicts the cumulative percentages for the papers published in 2009 and earlier while Figure 2(b) depicts the same for papers published in 2010 and later.

In Figure 2(a) and (b), "flat" periods—where cumulative percentage does not change very much over a few years—generally indicates that patents filed during that time were not directed toward the same area of technology that the paper is in. For example, for the Translation Caching paper, the cumulative percentage only increased from 21.5% to 27.8% between 2013 and 2016. This may indicate that patents in that timeframe were not directed toward the translation lookaside buffer (TLB), or perhaps toward caches in general. On the other hand, a sharp increase in the cumulative

percentage over a short period of time could indicate an increasing number of patents directed towards the paper's area of technology or that the ideas disclosed in the paper may finally be relevant to companies. Alternatively, a sharp increase may indicate that companies in that area of technology started filing significantly more patents or started citing that paper more consistently. For example, the cumulative percentage for the Translation Cache paper increased from 27.8% in 2016 to 73.4% in 2018. Similarly, the cumulative percentage for the DianNao paper increased from 31.6% in 2017 to 96.9% in 2018. These sharp increases in the cumulative percentages may indicate that TLBs and machine learning became the "hot" area of patented inventions during those respective timeframes.

REFERENCES

1. J. J. Yi, "Analysis of historical patenting behavior and patent characteristics of computer architecture companies—part V: References," *IEEE Micro*, Nov./Dec. 2022, Art. no. 135.
2. L. Branstetter and Y. Ogura, "Is academic science driving a surge in industrial innovation? Evidence from patent citations," Nat. Bureau Economic Res., Cambridge, MA, USA, Working Paper 11561, 2005.

JOSHUA J. YI is a solo practitioner who serves as a court appointed technical advisor for the honorable Alan D Albright, United States District Judge for the Western District of Texas, Waco Division, Waco, TX, USA. His research interests include microarchitecture and performance methodology. Yi received a Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA. Contact him at josh@joshuayipatentlaw.com.