UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

BRIGHT DATA LTD.,

Petitioner

v.

OXYLABS, UAB,

Patent Owner

IPR2023-00840

U.S. Patent No. 10,965,770

PETITION FOR *INTER PARTES* REVIEW

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

PATENT OWNER'S LIST OF EXHIBITS

Exhibit Number	Description
1001	U.S. Patent No. 10,965,770
1002	File Wrapper of U.S. Patent No. 10,965,770
1003	Declaration of Dr. Tim Arthur Williams
1004	International Application Publication Number WO 01/50320 to Cousins et al.
1005	Network Working Group Request for Comments: 7231 (2014)
1006	U.S. Patent Publication No. 2017/0180382 to Taylor et al.
1007	U.S. Patent No. 10,348,687 to Taylor et al.
1008	Network Working Group Request for Comments: 7540 (2015)
1009	Wessels, "Squid: The Definitive Guide," ISBN-10: 9780596001629, ISBN-13: 978-0596001629, O'Reilly Media; 1st Ed. (January 1, 2004)
1010	Squid Website at http://web.archive.org/web/20180920160052/https://wiki.squid-cache.org/ConfigExamples/Strange/RotatingIPs (accessed 3/30/2023)
1011	Appendix of Challenged Claims

TABLE OF AUTHORITIES

Cases

Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH, IPR2019-01469, Paper 6 at 7-9 (PTAB Feb. 13, 2020) 7
Bright Data Ltd. f/k/a Luminati Networks Ltd. v. Code200, UAB, et al., No. 2:19-cv-396 (E.D. Tex.) 4
Bright Data Ltd. f/k/a Luminati Networks Ltd. v. Teso LT, UAB, et al., No. 2:19-cv-395 (E.D. Tex.) 4
General Plastic Co., Ltd. v. Canon Kabushiki Kaisha, IPR2016-01357, Paper 19 at 15-16 (PTAB Sept. 6, 2017) 7
Luminati Networks Ltd. v. UAB Tesonet, No. 2:18-cv-299 (E.D. Tex.).....4
Teso LT, UAB v. Luminati Networks Ltd., No. 2:20-cv-00073 (E.D. Tex.) 5

Statutes

35 U.S.C. § 314(a) 1

Rules

37 C.F.R. § 42.103 6
37 C.F.R. § 42.104(a).....6
37 C.F.R. § 42.104(b)(1)-(2).....6
37 C.F.R. § 42.108(c).....1
37 C.F.R. § 42.15(a).....6
37 C.F.R. § 42.8 1

TABLE OF CONTENTS

I. INTRODUCTION 1

II. STATUTORY PREDICATES 1

A. Mandatory Notices (37 C.F.R. §42.8) 1

1. Real Parties-in-Interest 1
2. Related Matters 1
3. Lead and Back-Up Counsel 5
4. Service Information 5

B. Payment of Fees (37 C.F.R. §42.103) 6

C. Certification of Standing (37 C.F.R. §42.104(a)) 6

D. Identification of Challenges (37 C.F.R. §42.104(b)(1)-(2)) 6

III. DISCRETIONARY FACTORS 7

IV. OVERVIEW OF THE '770 PATENT 8

A. Summary of the '770 Patent 8

B. Claims of the '770 Patent 12

C. Prosecution History of the '770 Patent 12

D. Level of Skill in the Art 13

E. Claim Construction 13

V. OVERVIEW OF PRIOR ART REFERENCES 14

A. Cousins (EX.1004) 14

B. RFC7231 (EX.1005) 16

C. Taylor (EX.1006) 17

D. RFC7540 (EX.1008) 20

E. SQUID – THE DEFINITIVE GUIDE (EX.1009) 20

F. SQUID – WEBSITE (EX.1010) 21

VI. GROUNDS FOR INVALIDITY 21

A. Ground 1: Anticipation of Claims 1, 4-6, 9-14, 16, 17, 19, and 22-26 by Cousins....21
1. Claim 1....22
2. Claim 16....31
3. Claim 19....37
4. Claims 4 and 22....39
5. Claims 5 and 23....40
6. Claim 6....40
7. Claim 9....41
8. Claims 10 and 26....42
9. Claim 11....43
10. Claim 12....44
11. Claim 13....45
12. Claim 14....45
13. Claim 17....46
14. Claim 24....47
15. Claim 25....48

B. Ground 2: Obviousness of Claims 1, 2, 3, and 16-21 by Cousins in View of the Common Knowledge of a POSA....49
1. Claim 1....49
2. Claim 16....51
3. Claim 19....53
4. Claim 2....54
5. Claim 3....56
6. Claim 17....58
7. Claim 18....58
8. Claim 20....59
9. Claim 21....59

C. Ground 3: Obviousness of Claims 1, 7, 8, 13, and 16-19 by Cousins in View of Taylor....60
1. The Combination of Cousins and Taylor....60
2. Claim 1....62
3. Claim 16....63
4. Claim 19....66
5. Claim 7....68
6. Claim 8....69

7. Claim 13 70
8. Claim 15 71
9. Claim 17 72
10. Claim 18 72

D. Ground 4: Obviousness of Claims 2, 3, 20, and 21 Based on Cousins in View of RFC7231....72
1. The Combination of Cousins and RFC7231 72
2. Claims 2 and 20....74
3. Claims 3 and 21....75

E. Ground 5: Obviousness of Claim 18 Based on Cousins in View of RFC7540 76
1. The Combination of Cousins and RFC7540 76
1. Claim 18 78

VII. CONCLUSION 79

I. INTRODUCTION

Petitioner Bright Data Ltd. respectfully requests *inter partes* review and cancellation of claims 1-26 ("Challenged Claims") of Patent No. 10,965,770 (EX.1001, the "'770 Patent"). As outlined in this petition, there is a reasonable likelihood that at least one of the Challenged Claims is unpatentable. 35 U.S.C. §314(a); 37 C.F.R. §42.108(c). This petition is supported by the expert declaration of Dr. Tim A. Williams (EX.1003).

II. STATUTORY PREDICATES

A. Mandatory Notices (37 C.F.R. §42.8)

Pursuant to 37 C.F.R. §42.8, Petitioner Bright Data Ltd. hereby provides the following mandatory notices in connection with this *inter partes* review of Patent No. 10,965,770 titled "Dynamic Optimization of Request Parameters for Proxy Server."

1. Real Parties-in-Interest

Petitioner certifies that Bright Data Ltd. is the real party-in-interest.

2. Related Matters

There are no related district court cases involving Patent No. 10,965,770.

There are other matters between Petitioner and Patent Owner that do not involve Patent No. 10,965,770.

Petitioner has previously filed for inter partes review in the following:

- IPR2021-00983 (institution denied), challenging Patent No. 10,601,948;
- IPR2023-00185 (institution pending), challenging Patent No. 9,516,091;
- IPR2023-00186 (institution pending), challenging Patent No. 9,503,498; and
- IPR2023-00187 (institution pending), challenging Patent No. 8,924,457.

Patent Owner and/or one or more of its affiliates have filed for *inter partes* review of patents assigned to Petitioner Bright Data Ltd. They are:

- IPR2020-01266 (institution denied);
- IPR2020-01358 (institution denied);
- IPR2020-01506 (institution denied);
- IPR2021-00122 (institution denied);
- IPR2021-00249 (institution denied);
- IPR2021-01502 (institution denied);
- IPR2021-01503 (institution denied);
- IPR2022-00103 (pending);
- IPR2022-00353 (pending);

- IPR2022-00687 (pending);
- IPR2022-00861 (terminated);
- IPR2022-00862 (terminated);
- IPR2022-00936 (institution denied);
- IPR2022-01109 (institution denied);
- IPR2022-01110 (institution denied);
- IPR2023-00038 (terminated);
- IPR2023-00039 (terminated);
- IPR2021-01492 (pending); and
- IPR2021-01493 (pending).

Patent Owner and/or one or more of its affiliates have filed for *post grant* review of patents assigned to Petitioner Bright Data Ltd. They are:

- PGR2022-00052 of Patent No. 11,272,034 (pending); and
- PGR2022-00061 of Patent No. 11,349,953 (pending).

Patent Owner and/or one or more of its affiliates have filed *Ex Parte* Requests for Reexamination with the USPTO on patents assigned to Petitioner. They are:

- Control No. 90/014,816 (pending);

- Control No. 90/014,827 (pending);
- Control No. 90/014,875 (pending; currently stayed pending termination/completion of IPR2021-01492);
- Control No. 90/014,876 (pending; currently stayed pending termination/completion of IPR2021-01493); and
- Control No. 90/019,025 (denied).

There are other district court cases between Petitioner and Patent Owner and/or one or more of its affiliates that involve patents assigned to Petitioner and/or Patent Owner. They are styled as follows:

- *Luminati Networks Ltd. v. UAB Tesonet*, No. 2:18-cv-299 (E.D. Tex.)(closed);
- *Bright Data Ltd. f/k/a Luminati Networks Ltd. v. Teso LT, UAB, et al.*, No. 2:19-cv-395 (E.D. Tex.)(jury verdict entered; asserted patents found valid and infringed);
- *Bright Data Ltd. f/k/a Luminati Networks Ltd. v. Code200, UAB, et al.*, No. 2:19-cv-396 (E.D. Tex.)(pending);
- *Metacluster LT, UAB v. Bright Data Ltd.*, No. 2:22-cv-11 (E.D. Tex.)(pending); and
- *Bright Data Ltd. v. Oxylabs, UAB*, No. 2:23-cv-171 (E.D. Tex.)(pending).

There was another district court case between Petitioner and its affiliate Teso LT, UAB related to business torts, styled as follows: *Teso LT, UAB v. Luminati Networks Ltd.,* No. 2:20-cv-00073 (E.D. Tex.)(closed).

3. Lead and Back-Up Counsel

Lead Counsel	Back-Up Counsel
Thomas M. Dunham Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 838-1567 Fax (202) 478-1715 tomd@cherianllp.com USPTO Customer No. 144371 USPTO Reg. No. 39,965	Robert W. Molitors Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 838-1563 Fax (202) 478-1715 rmolitors@cherianllp.com USPTO Customer No. 144371 USPTO Reg. No. 66,726 Elizabeth A. O'Brien Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 873-1726 Fax (202) 478-1715 elizabetho@cherianllp.com USPTO Customer No. 144371 USPTO Reg. No. 76,105

4. Service Information

Electronic Mail	1. tomd@cherianllp.com (for correspondence from Patent Owner) or tom@dunham.cc (for correspondence from P-TACTS) 2. rmolitors@cherianllp.com

	3. elizabetho@cherianllp.com
Postal (and hand delivery) mailing addresses:	Thomas M. Dunham Robert W. Molitors Elizabeth A. O'Brien Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036
Telephone	Tel. (202) 838-1567
Facsimile:	Fax (202) 478-1715

Additionally, Petitioner consents to electronic service via email at the email addresses noted above.

B. Payment of Fees (37 C.F.R. §42.103)

The Office is authorized to charge Deposit Account No. 603803 (Customer No. 144371) the fees set forth in 37 C.F.R. §42.15(a), or any other applicable fees, for this Petition.

C. Certification of Standing (37 C.F.R. §42.104(a))

Petitioner certifies that Patent No. 10,965,770 is available for *inter partes* review and that Petitioner is not barred or estopped from requesting *inter partes* review of the Challenged Claims on the grounds alleged herein.

D. Identification of Challenges (37 C.F.R. §42.104(b)(1)-(2))

Petitioner requests cancellation of the Challenged Claims on the following grounds (independent claims underlined):

Ground 1	Claims 1, 4-6, 9-14, 16, 17, 19, and 22-26	102 Cousins
Ground 2	Claims 1, 2, 3, 16, 17, 18, 19, 20, and 21	103 Cousins + POSA's knowledge
Ground 3	Claims 1, 7, 8, 13, 16, 17, 18, and 19	103 Cousins + Taylor
Ground 4	Claims 2, 3, 20, and 21	103 Cousins + RFC7231
Ground 5	Claim 18	103 Cousins + RFC7540

III. DISCRETIONARY FACTORS

The factors considered under 35 U.S.C. §§314 and 325 do not weigh in favor of exercising discretion to deny institution.

The '770 Patent has not been challenged in any prior administrative proceeding. Thus, none of the factors set forth in *General Plastic* apply to this petition. *See General Plastic Co., Ltd. v. Canon Kabushiki Kaisha*, IPR2016-01357, Paper 19 at 15-16 (PTAB Sept. 6, 2017).

The prior art presented in this petition is not the same or substantially the same as the art previously presented during prosecution of the '770 Patent. As such, none of the factors set forth in *Advanced Bionics* apply to this Petition. *See Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 7-9 (PTAB Feb. 13, 2020).

Lastly, there is no parallel district court litigation regarding the '770 Patent.

IV. OVERVIEW OF THE '770 PATENT

U.S. Patent No. 10,965,770 ("the '770 Patent," EX.1001) issued from U.S. Application No. 17/018,333 filed on 9/11/2020. EX.1001, codes (21) and (22).

A. Summary of the '770 Patent

The '770 patent relates to the management of web crawling using proxies executing scraping strategies. Fig. 1 of the '770 Patent (reproduced below) shows a system with a service provider infrastructure 114 positioned between a user device 102 and target servers 116 and 118. EX.1001, 8:59-9:16; EX.1003, ¶32.

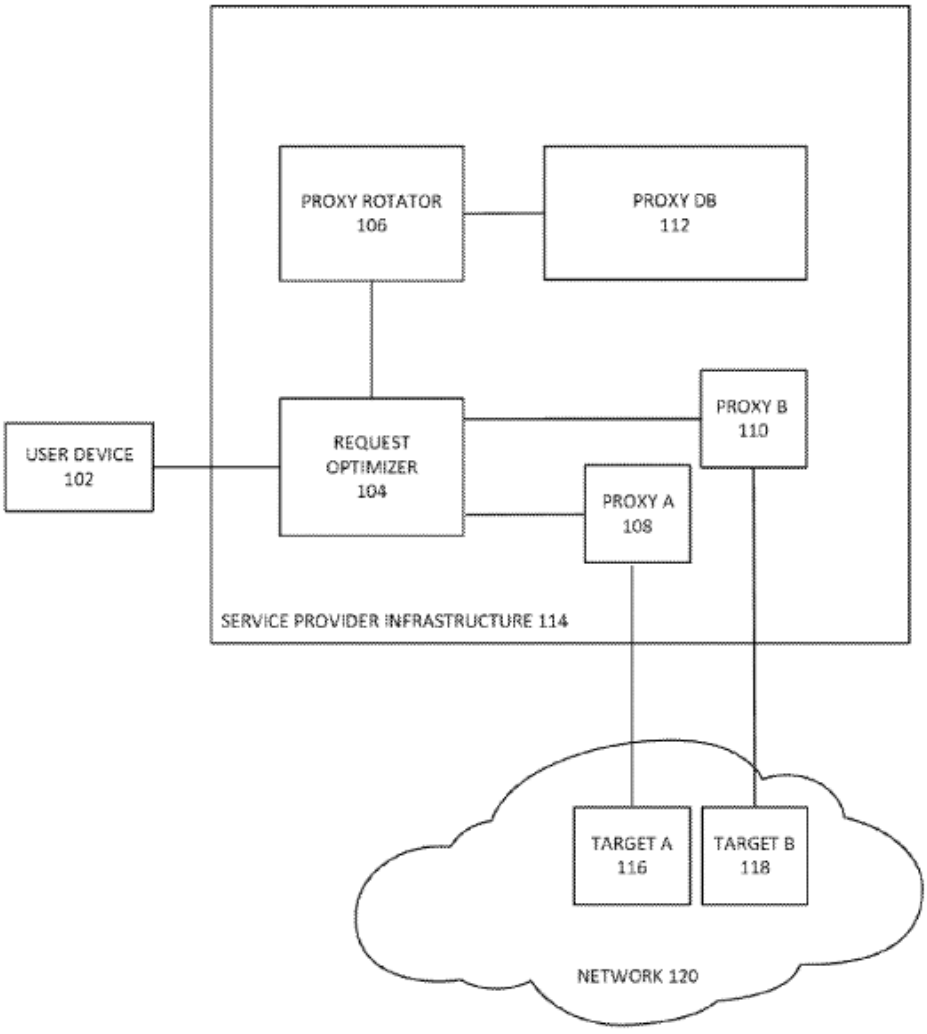


FIGURE 1

The user device is disclosed as being outside of the service provider infrastructure implementing a browser which may communicate a hypertext transfer protocol (HTTP) request for a web page stored on target 116 or target 118. EX.1001, 8:59-9:16; EX.1003, ¶32. Request optimizer 104 accepts and implements requests from the User Device. The Request Optimizer can read request metadata and adjust

the request using different pre-established Scraping strategies. EX.1001, 7:12-23; EX.1003, ¶32.

The parameters of a strategy may be adjusted if a negative reply received from the target, where the parameters may include a URL address, a header, cookies, or the body of the request. EX.1001, 17:10-20; EX.1003, ¶¶33-35.

Adjusting may include a modification of the query received from the user device. EX.1001, 10:34-36; EX.1003, ¶36. The request optimizer may also adjust parameters to increase efficiency such that if a URL address is incorrect the request optimizer may correct the URL address. EX.1001, 11:17-23; EX.1003, ¶36. FIG. 5 of the ’770 patent (reproduced below) illustrates the request being received by the request optimizer that includes a variety of scraping strategies that may be used to adjust the request.

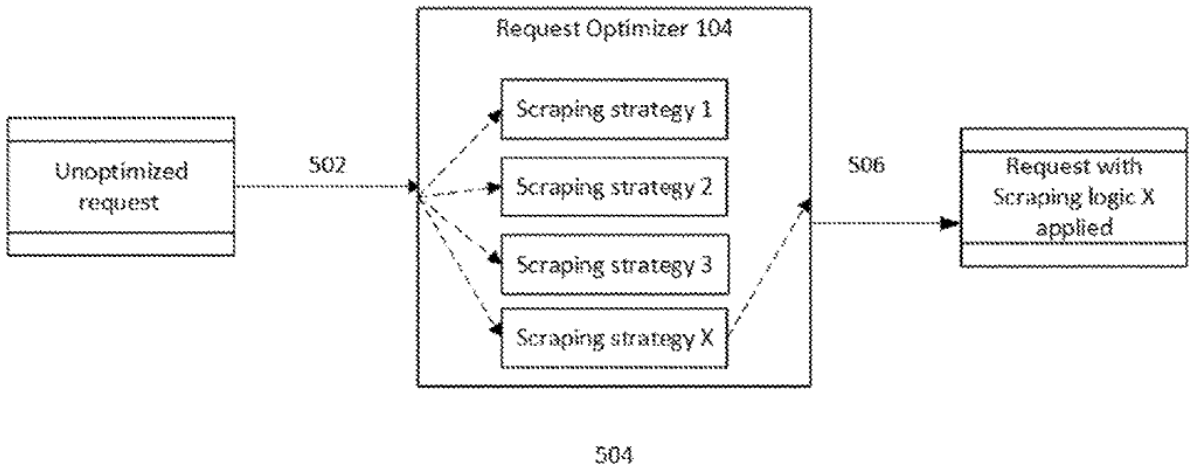


Figure 5

The request optimizer includes a scraping strategy where the parameters or policies may be optimized for a particular type of request or for a particular target. EX.1001, 8:27-29; EX.1003, ¶37. Scraping strategies may be developed by analyzing a target server's parameters and the target server's response data. EX.1001, 9:21-25; EX.1003, ¶37. Strategies may include selecting a proxy based on a geographic location. EX.1001, 9:43-44; EX.1003, ¶37.

Once the request parameters are adjusted the request optimizer may ask the proxy rotator for a particular proxy that is suitable to implement the user's request. EX.1001, 12:30-36; EX.1003, ¶38. The '770 patent discloses the sending of the adjusted request to a target that includes selecting the proxy according to the scraping strategy. EX.1001, 18:2-4; EX.1003, ¶38.

B. Claims of the ’770 Patent

The ’770 Patent includes 26 claims, with claims 1, 16, and 19 being independent. This Petition challenges claims 1-26. A Claims Appendix is provided in EX.1011.

C. Prosecution History of the ’770 Patent

The underlying application for the ’770 patent was filed on September 11, 2020, with 26 claims. EX.1002, 94-97. Applicant concurrently submitted a Track One request, which was granted on October 8, 2020. EX.1003, ¶39.

The United States Patent and Trademark Office (“Office”) issued an Office action on December 9, 2020 objecting to the claims due to various informalities but indicated that all claims would be allowable if rewritten or amended to overcome the informalities objections. EX.1003, ¶40.

On December 28, 2020 Applicant submitted a response amending claims to overcome the informalities objections. EX.1003, ¶41.

On December 31, 2020 the Office held an interview with Applicant to discuss amendments to claim 18 to overcome an informalities objection. The Office issued a Notice of Allowance on January 12, 2021. EX.1003, ¶¶42-43. The ’770 Patent issued on March 30, 2021. EX.1003, ¶44.

D. Level of Skill in the Art

The relevant technology for the ’770 Patent pertains to data retrieval of web sites using crawlers across networks. A person of ordinary skill in the art (“POSA”) would have had, as of the filing date of the ’770 patent, either: a Master’s Degree or higher in the field of electrical engineering, computer engineering, or computer science with a knowledge of Internet Communications and standards, or as of that time had a Bachelor’s Degree in the same fields and two or more years of experience in Internet Communications and standards. *See* EX.1003, ¶¶30-31.

E. Claim Construction

The claim terms at issue in the Challenged Claims have not been construed by any court.

The ’770 patent states that a proxy rotator “is a part of the Service Provider Infrastructure and is coupled with separate external components implementing specific assisting functionalities and is responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting.” EX.1001, 7:24-28. A POSA would understand “proxy rotator” in the context of the ’770 Patent means “a system coupled with separate external components implementing a specific functionality

and responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting." EX.1003, ¶19.[1]

Petitioner gives the remaining claim terms their plain and ordinary meaning in light of the specification and prosecution history as would have been read by a POSA.[2]

V. OVERVIEW OF PRIOR ART REFERENCES

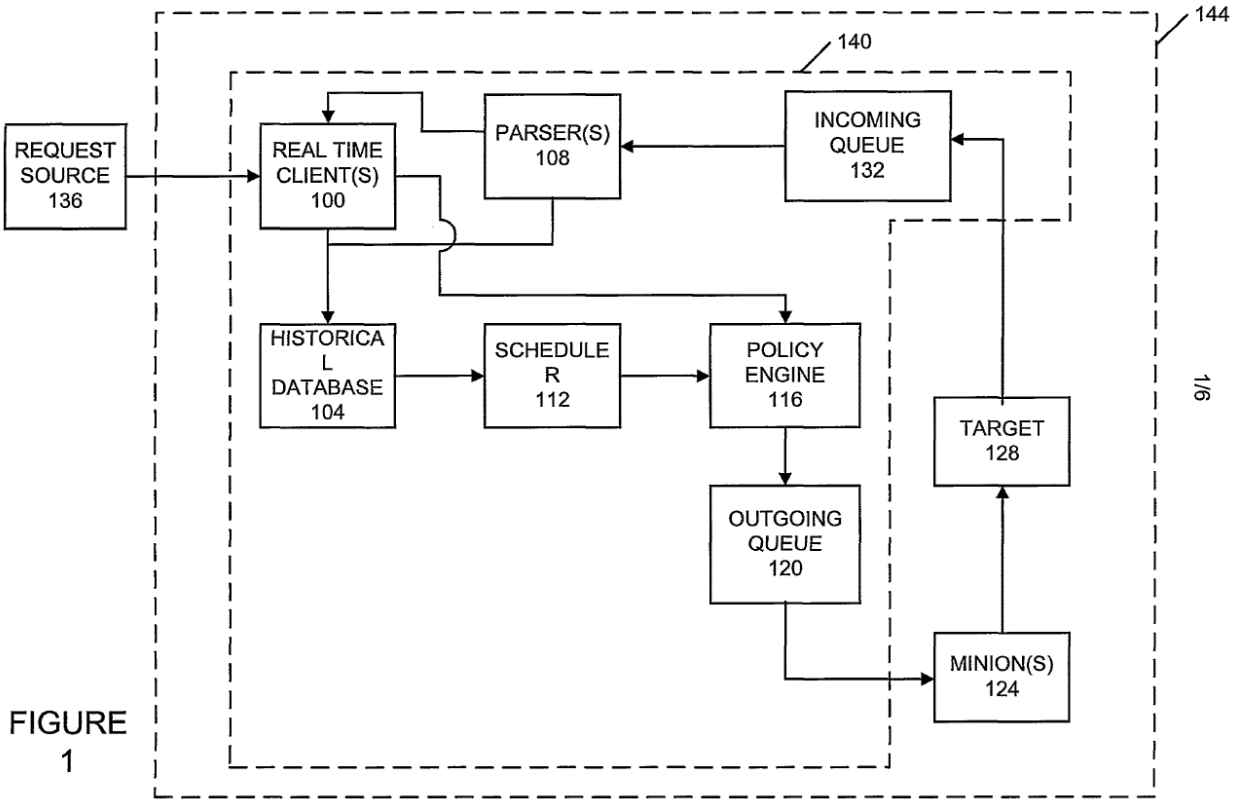
A. Cousins (EX.1004)

International Publication Number WO 2001/50320 to Cousins, et al. ("Cousins," EX.1004) was published on July 12, 2001, more than nineteen years before the September 11, 2020, filing date of the '770 Patent and qualifies as prior art under 35 U.S.C. §102(a)(1). Cousins is analogous to the '770 Patent at least because Cousins and the '770 Patent relate to the crawling and scaping of information on the web using a distributed network. EX.1003, ¶¶57-58.

Cousins describes a minimal impact crawler for searching and retrieving information based on policies for a target site. EX.1003, ¶59. The crawler employs

[1] Because the prior art discloses/teaches the narrower construction set forth above, the analysis herein does not change under a broader construction of the term "proxy rotator" consistent with the plain and ordinary meaning of the term.
[2] Petitioner reserves the right to submit proposed claim constructions if construed by a court and/or the Board.

a plurality of minions to retrieve crawl requests and crawl target web sites. *See* EX.1004, Abstract, code (57); EX.1003, ¶59. *See* also Fig. 1 below.



Cousins discloses the use of a policy engine that receives a user request and assembles policies for the target site. EX.1004, 3:13-15; EX.1003, ¶60. The requests may be ordered by priority to optimize data gathering in which the policy engine examines the crawl request. EX.1004, 4:17-22; EX.1003, ¶60. Cousins discloses an efficiency scraping policy to consolidate duplicate requests to the same target, where "two client engines 100 … could desire the same URL, but desire different information from the URL" where "the policy engine 116 combines the multiple

requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1994 at 11:18-24; EX.1003, ¶60.

Cousins also discloses a set of policies that include "Do not crawl …, Only crawl if …, Crawl anytime …, Never crawl." EX.1004, 10:25-32; EX.1003, ¶61.

Cousins discloses the use of minions, *e.g.*, proxies, where "[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data." EX.1004, 4:26-27; EX.1003, ¶62. The minions may be "located topologically as close as possible to the target site's server." EX.1004, 15:1-2; EX.1003, ¶62. Furthermore, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies." EX.1004, 16:28-29; EX.1003, ¶62.

In addition, "[b]y providing unique minions 124 to crawl web sites for which they are custom designed, the crawls can be much more efficient than a generic crawl." EX.1004, 16:10-14; EX.1003, ¶63.

B. RFC7231 (EX.1005)

Request for Comments 7231, titled "Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content," was published by the Internet Engineering Task Force ("IETF") Standards Track in June 2014. ("RFC7231," EX.1005). RFC7231 qualifies as a printed publication that was publicly accessible before the September 11, 2020,

filing date of the '770 Patent. *See* EX.1001, code (22); EX.1003, ¶64. RFC7231 qualifies as prior art under 35 U.S.C. §102(a)(1).

The IETF is an international non-profit standards organization that develops technical standards for Internet protocols. EX.1003, ¶65. A POSA would be very familiar with RFC7231 and the Internet standards explained within. *Id.*

RFC7231 was approved for publication by the Internet Engineering Steering Group of the IETF. *Id,* ¶66. RFC7231 is freely available for download at https://www.rfc-editor.org/rfc/rfc7231. *Id.* RFC7231 is intentionally disseminated per the IETF's open records policy. *Id.*

RFC7231 is analogous art to the '770 patent at least because RFC7231 and the '770 patent relate to accessing web pages utilizing the Hypertext Transfer Protocol. *Id,* ¶67. As discussed below, RFC7231 and the '770 patent involve using HTTP and metadata. *See, e.g.*, EX.1001, 3:11-22; 7:18-20; 7:48-50; 12:56-59; 13:60-67; EX.1003, ¶¶67-68.

C. Taylor (EX.1006)

U.S. Publication No. 2017/0180382 to Taylor et al. ("Taylor," EX.1006) was published on June 22, 2017, more than one year before the filing date of the '770

Patent.[3] EX.1001, code (22). Taylor qualifies as prior art under 35 U.S.C. §102(a)(1). EX.1003, ¶69.

Taylor is analogous art to the '770 Patent at least because Taylor and the '770 Patent relate to management of proxy servers within a service provider infrastructure. EX.1003, ¶70. Taylor and the '770 Patent relate to the transmission of digital information as Taylor and the '770 Patent are similarly classified in Intl. Cl. "H04L 29" transmission of digital information relating to arrangements, apparatus, circuits, or systems. EX.1001, code (51); EX.1006, code (51); EX.1007, code (51); EX.1003, ¶70. *See* also Fig. 1 reproduced below.

[3] Taylor issued as U.S. Patent No. 10,348,687 on July 9, 2019. EX.1007.

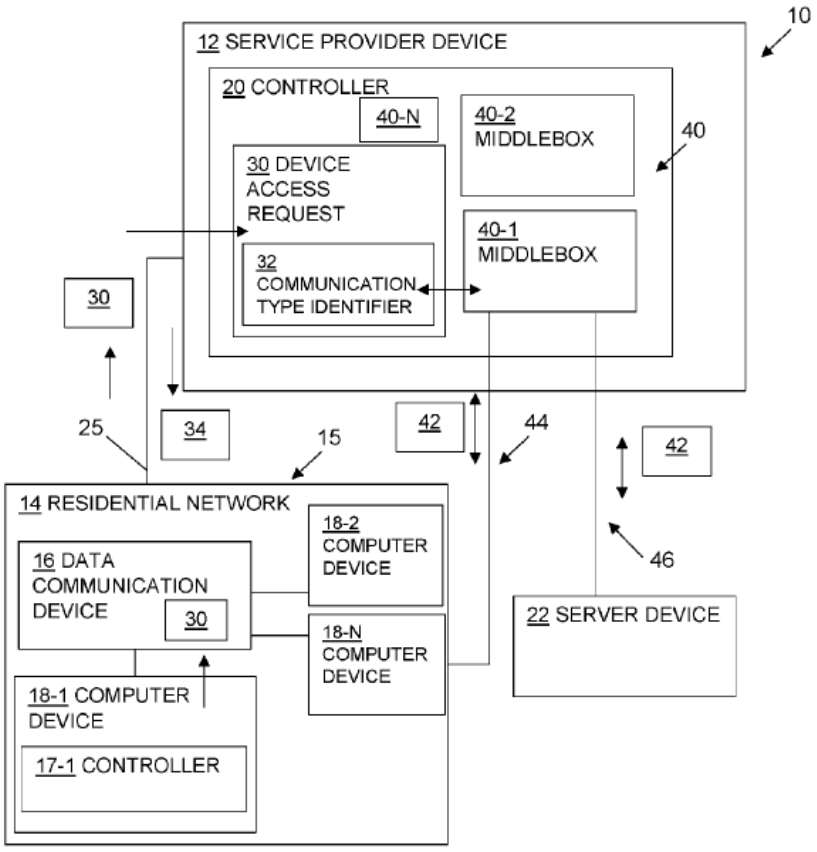


FIG. 1

Taylor is titled "Method and Apparatus for Using Software Defined Networking and Network Function Virtualization to Secure Residential Networks." EX.1006, code (54). Taylor discloses a service provider device that includes a controller configured to receive a request from a client device and establishing a secure communication channel between the client device and a middlebox. EX.1006, Abstract, code (57); EX.1003, ¶71. Taylor discloses selecting a middlebox proxy based on a correspondence of the destination IP address and a domain name server database. EX.1006, [0044]; EX.1003, ¶71.

D. RFC7540 (EX.1008)

Request for Comments 7540, titled "Hypertext Transfer Protocol – HTTP/2" was published by the IETF Standards Track in May 2015. ("RFC7540," EX.1008). RFC7540 qualifies as a printed publication that was publicly accessible before the September 11, 2020, filing date of the '770 Patent. *See* EX.1001, code (22); EX.1003, ¶72. RFC7540 qualifies as prior art under 35 U.S.C. §102(a)(1). EX.1003, ¶72.

EX.1003,RFC7540 was approved for publication by the Internet Engineering Steering Group of the IETF. RFC7540 is freely available for download at https://www.rfc-editor.org/rfc/rfc7540. *Id,* ¶74. RFC7540 is intentionally disseminated per the IETF's open records policy. *Id.*

RFC7540 is analogous art to the '770 patent at least because RFC7540 and the '770 patent relate to accessing web pages utilizing the Hypertext Transfer Protocol. *Id,* ¶75. As discussed below, RFC7540 and the '770 patent involve using HTTP and maintaining connections. *See, e.g.*, EX.1001, 3:11-22; 3:32-34; 15:19-22; 15:36-38; EX.1003, ¶¶75-76.

E. SQUID – THE DEFINITIVE GUIDE (EX.1009)

The open-source Squid proxy is documented in Wessels, EX.1009 hereto ("Squid"). The book was published by a major publisher in the field and has the

usual conventional markers (copyright, ISBN, etc.) indicating a reliable publication date in 2004. EX.1003, ¶77.

Squid is an "open source caching proxy for the Web." EX.1009, 1; EX.1003, ¶78. Squid discloses "both a proxy and a cache" where "[a]s a proxy, Squid is an intermediary in a web transaction" where "[i]t accepts a request from a client, processes that request, and then forwards the request to the origin server." EX.1009, 1; EX.1003, ¶78. And, as "a cache, Squid stores recently retrieved web content for possible reuse later" and where "[s]ubsequent requests for the same content may be served from the cache, rather than contacting the origin server again." EX.1009, 1; EX.1003, ¶78.

F. SQUID – WEBSITE (EX.1010)

The Squid website from a web.archive.org capture on September 20, 2018 discusses "Rotating the Squid outbounds IPs" and presents sample code for "Splitting uploads and downloads between two ISP links." EX.1010, 1; EX.1003, ¶79.

VI. GROUNDS FOR INVALIDITY

A. Ground 1: Anticipation of Claims 1, 4-6, 9-14, 16, 17, 19, and 22-26 by Cousins

1. Claim 1

a. [Claim 1a] "A method for data extraction and gathering from a web server comprising: …"

To the extent the preamble is limiting, Cousins discloses this claim limitation. EX.1003, ¶80. Cousins discloses "[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network." EX.1004, Abstract; EX.1003, ¶81. Fig. 1 of Cousins shows a web server, *i.e.*, target 128, where "[t]he crawler (144) applies the policies to schedule crawls (144) of the target site (128) and stores data retrieved from the crawl (144)." EX.1004, Abstract; EX.1003, ¶81.

Further, "[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention." EX.1004, 6:8-10; EX.1003, ¶82. "As crawlers are computer-driven, they can access large amounts of data very quickly, and they can perform crawls continuously." EX.1004, 1:20-21; EX.1003, ¶82.

b. [Claim 1b] "… comparing parameters of a request received from a user device with a list of scraping strategies …"

Cousins discloses "[a] request source [that] is any computer process that requests information from a data address" where "a user will access the crawler's web site through the user's Internet connection, and then input a request for a specific

item into a form provided by the crawler's web site" EX.1004, 6:15-24. Cousins discloses "scheduled" and "user-demanded crawl[s]." EX.1004, 3:26; EX.1003, ¶83. According to Dr. Williams, "Cousins disclosure of a user's access through an Internet connection, inputting a request for a specific item into a form, as well as the creation of a user-demanded crawl discloses a user request from a user device." EX.1003, ¶83.

Cousins states that the parameters of a request include a request for a specific item, for example, "a policy engine is provided that receives a request for a specific item and assembles polices for the target site containing information about the specific item." EX.1004, 3:13-15; EX.1003, ¶84. Cousins discloses comparing the request parameters with lists of policies, *e.g.*, scraping strategies, that includes where "[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc." EX.1004, 3:13-16; EX.1003, ¶84. Cousins discusses additional scraping strategies for a target site that may include best practices and/or efficiency based criteria such as "[d]o not crawl between 9:00 a.m. and 5:00 p.m." or "[d]o not crawl if web site is shut down", or "[n]ever crawl." EX.1004, 3:17-18; EX.1003, ¶84. Cousins discloses that such "policies allow the crawler to tailor its crawl to the requirements and preferences of the crawled site." EX.1004, 3:18-19; EX.1003, ¶84.

c. [Claim 1c] "… selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request;…"

Cousins describes one of the strategies as creating a more efficient scrape where "in a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL" and "prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128" such that "[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1004, 11:17-24; EX.1003, ¶85. Further, "by coalescing multiple requests from different request sources 136 for similar data into a single crawl, the crawler 144 of the present invention provides lower overhead, lower latency and globally higher throughput. It also decreases the crawling burden on the target site 128." EX.1004. at 11:27-30; EX.1003, ¶85.

Another strategy is time based so that the most current information is obtained, where "[f]or example, a policy for a target site 128 may be 'crawl if the data is older than ten hours'" where "[i]n this example, the client engine 100 passes a request to the policy engine 116, as discussed below, which then checks a policy stored for the target site 128 and the date/time stamp of the data in the historical

database 104 that would satisfy the request to determine whether the data in the database 104 may be used to satisfy the request." EX.1004, 8:16-21; EX.1003, ¶86.

Cousins also describes a general selecting of a scraping strategy where "a database may be created for time of day, or target site, or connectivity loss, or site closure" where the policy engine "will examine the time of day database to determine if a policy is applicable to the crawl request based on the current time of day" and that "[t]he policy engine 116 will examine all of the databases to build a list of all policies that apply to a particular request for data." EX.1004, 10:9-12; EX.1003, ¶87.

d. [Claim 1d] "… identifying discrepancies between the request and the scraping strategy;…"

The '770 patent does not define what may be classified as a discrepancy. As defined by the Merriam-Webster dictionary a discrepancy is "1: the quality or state of disagreeing or being at variance" and "2: an instance of disagreeing or being at variance."[4] EX.1003, ¶88. As stated by Dr. Williams, "a POSA would have understood identifying a discrepancy as identifying a disagreement or some difference between the parameters of the received request and the parameters of a request according to the selected scraping strategy." *Id.*

[4] https://www.merriam-webster.com/dictionary/discrepancy, accessed on April 11, 2023.

An example of a strategy including a discrepancy was discussed above, where there are two separate received requests where both requests are for the same target site. Instead of executing the two separate requests the strategy "combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1004, 11:17-24; EX.1003, ¶89.

Cousins describes another example discrepancy pertaining to the time when data is to be crawled from a target server where a request is made to retrieve "a current event news story on a web site," but a scraping strategy dictates that a recently executed crawl already has retrieved the requested data making a new crawl inefficient. EX.1003, ¶90. For example, "a request is for a current event news story … the window may require that the data have been crawled within the last four hours," but "[f]or data that changes less frequently, such as a description of an archived news story, the window may permit data to be retrieved from the database 104 if it has been crawled within the last seven days or longer" and "[i]n an embodiment in which the crawler 144 verifies the validity of data in the database 104, the policy engine 116 preferably stores the policies governing when a site 128 should be crawled." EX.1004, 8:10-16; EX.1003, ¶90.

Cousins explains that "item information and seller information do not change during the life of an auction" and "[t]herefore, the initial crawl can provide this

information, and then no further crawl is necessary unless a user request for price updates is received." EX.1004, 4:9-11; EX.1003, ¶91.

Cousins describes a discrepancy regarding the priority of competing requests. For example, two requests arrive both claiming priority, where "for example, if a first request to a target site 128 is determined to be appropriate for transmission, and is a user-demand request, and a second request to a different target site 128 is also determined to be acceptable but is from an automated discovery crawl, the user demand request is given a higher priority than the request from the automated discovery crawl." EX.1004, 12:21-26; EX.1003, ¶92.

e. [Claim 1e] "… adjusting the parameters of the request according to the scraping strategy to form an adjusted request;…"

Cousins describes adjusting a variety of request parameters to form an adjusted request. EX.1003, ¶93. Cousins also describes adjusting parameters by consolidating multiple requests as discussed above where "two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL" and "prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128" such that "[i]f there are, the policy engine 116 combines the multiple

requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1004, 11:17-24; EX.1003, ¶93.

Cousins also describes adjusting the use policy parameters of a request. For example, "[i]f a crawl request 400 is rescheduled by the scheduler 112, upon being received again by the policy engine 116, the policy engine 116 builds new policies for the rescheduled request to account for any new policies that may apply to the crawl request 400." EX.1004, 11:6-9; EX.1003, ¶94. In addition, "[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID[logical user id[5]] 300 to the policy engine 116 to generate policies for the request for data." EX.1004, 8:22-23; EX.1003, ¶94.

f. [Claim 1f] "… sending the adjusted request to a target through a proxy selected according to the scraping strategy; …"

Cousins discloses scraping strategies for a target site that include best practices and/or efficiency-based criteria. EX.1003, ¶95. Cousins discloses, due to network congestion or a network failure, the selection of a proxy is determined by

[5] "An LUID is a data structure used to store and organize data used by the crawler 144 and to uniquely identify each auction monitored by the crawler 144 of the present invention. ... The LUID also has a unique identification number 312 used by the crawler 144 to internally differentiate different auctions, or other collections of data. The use of LUIDs 300 to track data stored on target sites 128 provides a high degree of flexibility to the crawler 144." See page 7 at 1:12.

the scraping strategy where "a target site can be accessed from multiple different pathways" where "if one minion is unable to crawl a target site, because of congestion or network failures, a different minion located elsewhere in the Internet can still access the target site and provide the requested crawl." EX.1004, 4:33-5:3; EX.1003, ¶95. Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004. at 15:1-2 EX.1003, ¶95. In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004. at 16:28-29; EX.1003, ¶95. Squid discloses the uses of proxies, referred to as redirectors, that is "an external process that rewrites URIs from client requests" where "Squid fetches the new URI automatically, as though the client originally requested it." EX.1009, 185; EX.1003, ¶95.

g. [Claim 1g] "… receiving a reply from the target."

The applicant admitted prior art section in the '770 patent describes how the then-current proxy systems functioned, including that "a proxy operates in the following manner: i) receives a request from a user in a public network, or private network separated by a firewall, to a remote server that contains the desired content; ii) creates a proxy tunnel toward the remote server; iii) forwards the request to the remote server through the established tunnel; iv) receives the response from the

remote server; and v) forwards the response back to the user." EX.1001, 1:30-37; EX.1003, ¶96.

Cousins discloses receiving a reply from the target where "[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data." EX.1004, 4:26-27; EX.1003, ¶97. Also, Cousins describes the retrieval of data from the target where "[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention" and that "[a] crawl is defined as the act of fetching a collection of data, typically a web page, from the target site 128, or crawlee." EX.1004, 6:8-11; EX.1003, ¶97. Cousins also discloses that "policies are generated by an operator of the crawler 144 based on data received from the target site 128." EX.1004, 10:1-2; EX.1003, ¶97.

Cousins discloses receiving a reply from the target where "[t]he client engine 100 associates the LUID with a client engine identification and a parser identification" and "[t]his allows the crawler 144 to identify which parser 108 to use when the collection of data is retrieved from the target site 128, and to which client engine to route the requested data." EX.1004, 7:16-19; EX.1003, ¶98.

Thus, for the reasons discussed above, claim 1 is anticipated by Cousins. EX.1003, ¶99.

2. Claim 16

a. [Claim 16a] "A system for data extraction and gathering from a network comprising at least: …"

To the extent the preamble is limiting, Cousins discloses this claim limitation. EX.1003, ¶100. Fig. 1 of Cousins shows a web server, *i.e.*, target 128, and in Fig. 7, *i.e.*, target site server 708 in which "crawlers have been developed … that provide data gathering, searching, and retrieval services to allow users or Internet-related applications to locate desired data from the millions of web sites." EX.1004, 1:14-16; EX.1003, ¶100. Cousins discloses "[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network." EX.1004. at Abstract; EX.1003, ¶100.

Further, "[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention." EX.1004, 6:8-10; EX.1003, ¶101. "As crawlers are computer-driven, they can access large amounts of data very quickly, and they can perform crawls continuously." EX.1004. at 1:20-21; EX.1003, ¶101.

b. [Claim 16b] "… a proxy rotator operable to provide a plurality of proxies over the network; …"

As construed above, a "proxy rotator" is a system coupled with separate components implementing a specific functionality and responsible for proxy control,

rotation, maintenance, collecting statistical data, and reporting. EX.1003, ¶102. Cousins discloses a policy engine 116 that is responsible for the control of minions, *e.g.*, the proxies, where "[a] policy engine (116) is provided that receives a request for a specific item and assembles policies for the target site … [where] [t]he policies are rules that determine the crawl (144) of a target site (128) … [where] [t]he crawler (144) employs a plurality of minions (144) to retrieve crawl (144) requests and crawl (144) the target web sites (128) to obtain the necessary data." EX.1004, Abstract; EX.1003, 102. Further, "the policy engine 116 decides not only whether a request may be transmitted at that time for a given target site 128 but also decides the priority between competing requests for different target sites 128." EX.1004, 12:18-20; EX.1003, 102.

Cousins describes that the proxies, *e.g.*, minions, may be strategically located, for example, "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, 103. In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses." EX.1004, 16:28-29; EX.1003, 103. The policy engine may also rotate proxy requests through the use of multiple minions where "a target site can be accessed from multiple different pathways" where "a different minion located elsewhere in the Internet can still access the target site and provide the requested crawl."EX.1004, 4:33-5:3; EX.1003, 103.

Cousins also discloses where the policy engine can utilize multiple minions such that "a target site can be accessed from multiple different pathways … if one minion is unable to crawl a target site … a different minion located elsewhere in the Internet can still access the target site." EX.1004, 4:33-5:2; EX.1003, 104.

c. [Claim 16c] "… a request optimizer operable to compare parameters of a request received from a user device with a list of scraping strategies,…"

The '770 patent, in an embodiment, states that request optimizer 104 "is a next generation proxy that is a part of the service provider infrastructure that *accepts requests* from the User Device and *implements* the requests through the system." EX.1004, 7:12-15; EX.1003, ¶105. Using an analysis algorithm, the Request Optimizer can read the metadata of the request and adjust the request using different pre-established Scraping strategies to better scrape the Target. EX.1001, 7:12-23; EX.1003, 105.

Cousins discloses that the policy engine 116 *accepts requests and implements* the requests through the use of a crawler and minions where the "policy engine (116) is provided that receives a request for a specific item and assembles policies for the target site … [where] policies are rules … [where] the crawler (144) applies the policies to schedule crawls (144) … [where] [t]he crawler (144) employes a plurality of minions (144) to retrieve crawl (144) requests and crawl (144) the target web sites

(128) to obtain the necessary data." EX.1004, Abstract, emphasis added; EX.1003, 106.

Cousins discloses that the policy engine 116 receives a request where the "policy engine is provided that receives a request for a specific item and assembles polices for the target site containing information about the specific item." EX.1004, 3:13-15; EX.1003, 107.

Cousins discloses comparing the request parameters with lists of policies, *e.g.*, scraping strategies, that includes where "[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc." EX.1004, 3:13-16; EX.1003, ¶108. Cousins discusses additional possible scraping strategies for a target site that may include best practices and/or efficiency based criteria such as "[d]o not crawl between 9:00 a.m. and 5:00 p.m." or "[d]o not crawl if web site is shut down", or "[n]ever crawl." EX.1004, 3:17-18; EX.1003, ¶108. Cousins discloses that such "policies allow the crawler to tailor its crawl to the requirements and preferences of the crawled site." EX.1004, 3:18-19; EX.1003, ¶108.

d. [Claim 16d] "… to select a scraping strategy from the list of scraping strategies a scraping strategy that corresponds to the parameters of the request, …"

Petitioner incorporates by reference its discussion in Section VI.A.1.c herein. As discussed, Cousins discloses that the policy engine 116 "that receives a request for a specific item and assembles policies for the target site that contain information about the specific item." EX.1004, 3:13-15; EX.1003, ¶110.

e. [Claim 16e] "… to identify discrepancies between the request and the scraping strategy,…"

Petitioner incorporates by reference its discussion in Section VI.A.1.d herein.

f. [Claim 16f] "… to adjust the parameters of the request according to the scraping strategy to form an adjusted request, …"

Petitioner incorporates by reference its discussion in Section VI.A.1.e herein.

g. [Claim 16g] "… to send the adjusted request to a target through a proxy selected from a proxy database by the proxy rotator according to the scraping strategy."

Cousins discloses scraping strategies for a target site that include best practices and/or efficiency-based criteria. EX.1003, ¶113. Cousins discloses, due to network congestion or a network failure, the selection of a proxy is determined by the scraping strategy where "a target site can be accessed from multiple different pathways" where "if one minion is unable to crawl a target site, because of congestion or network failures, a different minion located elsewhere in the Internet

can still access the target site and provide the requested crawl." EX.1004, 4:33-5:3; EX.1003, ¶113. Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, ¶113. In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶113.

Cousins discloses that the data of a particular crawl of a selected target website is stored in a database, for example, "after a crawl, either scheduled or by request, the crawler stores data retrieved from the crawl into a historical database" and "when a scheduled or user-demanded crawl is to occur, the crawler first checks the historical database to determine whether the request can be satisfied from the data stored in the database." EX.1004, 3:24-26; EX.1003, ¶103. If the request cannot be satisfied from the stored data, then a new crawl is requested based on the minion selection in the database. EX.1004, 4:1-3; EX.1003, ¶114.

Thus, for the reasons discussed above, claim 16 is anticipated by Cousins. EX.1003, ¶115.

3. Claim 19

a. [Claim 19a] "…a non-transitory computer readable medium having computer readable program code embodied therewith, the computer readable program code comprising:…"

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Fig. 1 of Cousins shows a web server, *i.e.*, target 128, and in Fig. 7, *i.e.*, target site server 708 in which "crawlers have been developed … that provide data gathering, searching, and retrieval services to allow users or Internet-related applications to locate desired data from the millions of web sites." EX.1004, 1:14-16; EX.1003, ¶116. Cousins discloses "[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network." EX.1004, Abstract; EX.1003, ¶116.

Further, "[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention." EX.1004, 6:8-10; EX.1003, ¶117. "As crawlers are computer-driven, they can access large amounts of data very quickly, and they can perform crawls continuously." EX.1004, 1:20-21. Cousins discloses minions, where "[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data." EX.1004, 4:26-27; EX.1003, ¶117. Cousins discloses that "[m]inions 124 may also be implemented to be multithreaded

and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶117. Cousins discloses that "[i]n another embodiment, minions 124 use accounts at different Internet Service Providers." EX.1004, 16:27-28; EX.1003, ¶117.

b. [Claim 19b] "… comparing parameters of a request received from a user device with a list of scraping strategies;…"

Petitioner incorporates by reference its discussion in Section VI.A.1.b herein.

c. [Claim 19c] "… selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request,…"

Petitioner incorporates by reference its discussion in Section VI.A.1.c herein.

d. [Claim 19d] "… identifying discrepancies between the request and the scraping strategy, …"

Petitioner incorporates by reference its discussion in Section VI.A.1.d herein.

e. [Claim 19e] "… adjusting the parameters of the request according to the scraping strategy to form an adjusted request, …"

Petitioner incorporates by reference its discussion in Section VI.A.1.e herein.

f. [Claim 19f] "… sending the adjusted request to a target through a proxy selected according to the scraping strategy."

Petitioner incorporates by reference its discussion in Section VI.A.1.f herein. Thus, for the reasons discussed above, claim 19 is anticipated by Cousins. EX.1003, 123.

4. Claims 4 and 22

"… wherein if the reply received from the target is negative, the method repeats by adjusting the parameters of the request according to a different scraping strategy from the list of scraping strategies to form a different adjusted request that is then sent to the target."

Petitioner incorporates by reference its discussion in Section VI.A.1 and VI.A.3 herein.

Cousins discloses conducting a scrape, retrieving information from a target, and then storing that information in a database. However, "[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID 300 to the policy engine 116 to generate policies for the request for data," thus the new policies are used to form a new adjusted request for data from the target. EX.1004, 8:22-23; EX.1003, ¶125.

In another example, if the reply from the target is negative, for example, "if servers at the target site's main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number

of requests to the site 128," the parameters of the request may be further adjusted by retrieving an alternate exact URL such that "the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104" which "allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest." EX.1004, 8:33-9:5; EX.1003, ¶126. "Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning." EX.1004, 9:5-7; EX.1003, ¶¶126 and 141.

5. Claims 5 and 23

> *"... wherein the method is repeated with alternative scraping strategies selected from the list of scraping strategies until a positive response is obtained or until a maximum number of attempts is reached."*

Petitioner incorporates by reference its discussion in Sections VI.A.1, VI.A.3, and VI.A.4 herein.

As discussed in claims 4 and 22, Cousins discloses conducting an alternated scraping strategy to obtain a positive response. EX.1003, ¶¶127 and 142.

6. Claim 6

> *"... wherein at least one of the parameters within the following components of the request is adjusted: a. URL address; b. headers; c. cookies; d. request body."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

As discussed in claims 4 and 5, Cousins discloses adjusting a URL address. EX.1003, ¶128. For example, "if servers at the target site's main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128," the parameters of the request may be further adjusted by retrieving an alternate exact URL such that "the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104" which "allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest." EX.1004, 8:33-9:5; EX.1003, ¶128. "Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning." EX.1004, 9:5-7; EX.1003, ¶128.

7. Claim 9

> *"… wherein the proxy remains available for multiple user device requests."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Cousins discloses a queue of multiple user requests where "a minion 124 will pull the request from the queue 120 and fetch the page from the target site 128 that contains the requested data" and "[i]n a preferred embodiment, once a request is fed

to a minion 124, the output queue 120 automatically schedules a second crawl request 400 to be made in the near future on the expectation that the first crawl request 400 will fail." EX.1004, 8:13:4-8; EX.1003, ¶129.

Cousins discloses the management of multiple requests, for example, "[i]n a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL" and "[a]ccordingly, prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128" and "[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1004, 11:17-24; EX.1003, ¶130.

8. Claims 10 and 26

> *"… wherein the parameters of the request are not adjusted if the request received from the user device is aligned to one of the scraping strategies from the list."*

Petitioner incorporates by reference its discussion in Sections VI.A.1 and VI.A.3 herein.

Cousins discloses that a parameter of a request may include "a URI that identifies a specific item of data … [f]or example, a user may request the crawler

144 to track a particular auction for a camera that appears on an auction site." EX.1004, 6:26-28; EX.1003, ¶131. Cousins discloses that a scraping strategy based on a time stamp, for example, "if a request is for a current event news story … the window may require that the data have been crawled within the last four hours." EX.1004, 8:10-11; EX.1003, ¶131. Thus, if the stored crawl data is older than four hours, the requested crawl will proceed without any adjustment to the request parameters. EX.1003, ¶¶131 and 146.

9. Claim 11

> *"… wherein request optimizer operational settings include at least one of the following: scraping strategies, session ID numbers, servers accepted as request originators, operation timeframes, scheduled maintenance jobs, connectivity sessions timeouts, request processing logs, or any combination thereof and such settings are stored in a database on disk, in a flat file on disk, in a database in memory, or other storage media."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Cousins discloses this claim limitation. For example, scraping strategies, *e.g.*, policies, are stored in a database, such as "[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc." EX.1004, 3:15-16; EX.1003, ¶132.

Another example, Cousins discloses an operational setting including at least a session ID number, for example, "[i]n a preferred embodiment, the real time client

engine 100 converts 204 the request information into an LUID (logical user id)" and where "[a]n LUID is a data structure used to store and organize data used by the crawler 144 and to uniquely identify each auction monitored by the crawler 144 of the present invention." EX.1004, 6:33-7:3; EX.1003, ¶133. In addition, the "LUID also has a unique identification number 312." EX.1004, 7:9; EX.1003, ¶133.

Cousins discloses connectivity timeouts, as discussed above where "[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID 300 to the policy engine 116 to generate policies for the request for data," thus the new policies are used to form a new adjusted request for data from the target. EX.1004, 8:22-23; EX.1003, ¶134.

10. Claim 12

> *"… wherein the target is a web server serving the reply accessible by HTTP/HTTPS protocols and is attached to a network."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Cousins discloses this claim limitation. For example, "crawlers gather this information by using the Internet in a process known as crawling" where "[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet." EX.1004, 1:17-20; EX.1003, ¶135.

11. Claim 13

> *"… wherein the proxy is a proxy server integrated into a service provider infrastructure."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Cousins discloses this claim limitation. As discussed, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies." EX.1004, 16:28-29; EX.1003, ¶136.

Cousins discloses "in one embodiment, the crawler 144 collocates servers throughout the world, and minions 124 are stored on the collocated servers" and "[i]n another embodiment, minions 124 use accounts at different Internet Service Providers." EX.1004, 16:25-28; EX.1003, ¶137. Accordingly, Cousins discloses the use of minions and proxies in both scenarios, integrated into a service provider infrastructure and separate from a service provider infrastructure. EX.1003, ¶137.

12. Claim 14

> *"… wherein the proxy is a proxy server separate from a service provider infrastructure."*

Petitioner incorporates by reference its discussion in Section VI.A.1 herein.

Cousins discloses this claim limitation. As discussed, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies." EX.1004, 16:28-29; EX.1003, ¶138.

Cousins discloses "in one embodiment, the crawler 144 collocates servers throughout the world, and minions 124 are stored on the collocated servers" and "[i]n another embodiment, minions 124 use accounts at different Internet Service Providers." EX.1004, 16:25-28; EX.1003, ¶139. Accordingly, Cousins discloses the use of minions and proxies in both scenarios, integrated into a service provider infrastructure and separate from a service provider infrastructure. EX.1003, ¶139.

13. Claim 17

> *"… wherein the request optimizer can adjust the request received from the user device by modifying at least one of the parameters within the following components: a. URL address; b. headers; c. cookies; d. request body."*

Petitioner incorporates by reference its discussion in Section VI.A.2 herein.

Cousins discloses this claim limitation. For example, as discussed above in claims 4, 5, and 6, Cousins discloses adjusting a URL address. For example, "if servers at the target site's main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128," the parameters of the request may be further adjusted by retrieving an alternate exact URL such that "the crawler 144 of the present invention

is still able to satisfy requests because the exact URL for a given web page is stored in the database 104" which "allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest." EX.1004, 8:33-9:5; EX.1003, ¶140. "Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning." EX.1004, 9:5-7; EX.1003, ¶140.

14. Claim 24

> *"… wherein the proxy remains available for multiple user device requests."*

Petitioner incorporates by reference its discussion in Section VI.A.3 herein.

As discussed in claim 9, Cousins discloses a queue of multiple user requests where "a minion 124 will pull the request from the queue 120 and fetch the page from the target site 128 that contains the requested data" and "[i]n a preferred embodiment, once a request is fed to a minion 124, the output queue 120 automatically schedules a second crawl request 400 to be made in the near future on the expectation that the first crawl request 400 will fail." EX.1004, 8:13:4-8; EX.1003, ¶143.

Cousins discloses the management of multiple requests, for example, "[i]n a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from

the URL" and "[a]ccordingly, prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128" and "[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404." EX.1004, 11:17-24; EX.1003, ¶144.

15. Claim 25

> *"… wherein at least one of the following parameters of the request is adjusted: a. URL address; b. headers; c. cookies; d. request body."*

Petitioner incorporates by reference its discussion in Section VI.A.3, VI.A.6, and VI.A.13 herein.

As discussed in claims 4, 5, and 6, Cousins discloses adjusting a URL address. EX.1003, ¶145. For example, "if servers at the target site's main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128," the parameters of the request may be further adjusted by retrieving an alternate exact URL such that "the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104" which "allows the crawler 144 to bypass the main path to the target site 128 and access the target site's

direct address for the item of interest." EX.1004, 8:33-9:5; EX.1003, ¶145. "Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning." EX.1004, 9:5-7; EX.1003, ¶145.

B. Ground 2: Obviousness of Claims 1, 2, 3, and 16-21 by Cousins in View of the Common Knowledge of a POSA

Claims 1, 2, 3, and 16-21 of the '770 patent are obvious under Cousins in view of the common knowledge of a POSA. EX.1003, ¶147.

1. Claim 1

a. [Claim 1a], [Claim 1b], [Claim 1c], [Claim 1d], and [Claim 1e]

Petitioner incorporates by reference its discussion in Sections VI.A.1.a, VI.A.1.b, VI.A.1.c, VI.A.1.d, and VI.A.1.e herein.

b. [Claim 1f]

Petitioner incorporates by reference its discussion in Section VI.A.1.f herein.

If the Board does not agree that Cousins discloses "a proxy" then such a limitation would be common knowledge to a POSA. EX.1003, ¶¶153-154.

Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, ¶153.

In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶153. Further, Squid, as known by a POSA, is a proxy rotator. EX.1009, 1-2; EX.1003, ¶153.

It would have been obvious to a POSA to utilize minions as proxies or to have minions use proxies as Cousins specifically states such a configuration. EX.1003, ¶154. A POSA would also have known of the advantages of using proxies, for example as stated in applicant's own admitted prior art in the '770 patent, "[p]roxies can be used to bypass certain Internet restrictions (*e.g*. firewalls) by enabling a user to request the content through a (remote) proxy server instead of accessing the content directly" where "[p]roxy servers are often used to get around geo-IP based content restrictions." EX.1004, 2:25-29; EX.1003, ¶154. Dr. Williams concludes that it would have been obvious to combine Cousins with the knowledge of a POSA. EX.1003, ¶155.

c. [Claim 1g] "… receiving a reply from the target."

Petitioner incorporates by reference its discussion in Section VI.A.1.g herein.

Thus, for the reasons discussed above, claim 1 is obvious under Cousins in view of the common knowledge of a POSA. EX.1003, ¶157.

2. Claim 16

a. [Claim 16a], [Claim 16d], [Claim 16e], and [Claim 16f]

Petitioner incorporates by reference its discussion in Sections VI.A.2.a, VI.A.2.d, VI.A.2.e, and VI.A.2.f herein.

b. [Claim 16b]

Petitioner incorporates by reference its discussion in Section VI.A.2.b herein.

If the Board does not agree that Cousins discloses "a proxy rotator" as construed above, then such a limitation would be common knowledge to a POSA. EX.1003, ¶159.

Cousins explicitly states that the "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶149. Cousins explicit reference to the Squid proxy would be known by a person of ordinary skill in the art to be a proxy rotator as evidenced by the September 2018 web.archive capture of the Squid proxy website explaining the use of a proxy rotator entitled "Rotating the Squid outbound IPs." EX.1009, 1-2; EX.1003, ¶160. Thus, Squid's disclosure of "Rotating the Squid outbound IPs" would have been known by a POSA to a separate piece of software that could exist on any computing device, for example within policy engine 116, server 700, or outgoing queue 120. *Id.* Squid also discloses proxy rotating where

"Squid iterates between all addresses (a.k.a. round-robin) for each cache miss" and "to use a redirector (see Chapter 11) to select the backend server." EX.1009, 309; EX.1003, ¶160.

The '770 patent, in its background section, discusses the use of "rotating proxies" with a web crawler where "[t]he task of attempting data retrieval further includes managing Hypertext Transfer Protocol (HTTP) and Hypertext Transfer Protocol Secure (HTTPS) requests, rotating proxies, and headers." EX.1001, 3:7-10. Dr. Williams concludes it would have been obvious to adapt Cousins with the common knowledge of a POSA. EX.1003, ¶¶162-164.

c. [Claim 16c]

Petitioner incorporates by reference its discussion in Section VI.A.2.c herein.

The '770 patent, in an embodiment, states that request optimizer 104 "is a next generation proxy that is a part of the service provider infrastructure that accepts requests from the User Device and implements the requests through the system." EX.1004, 7:12-15. Cousins discloses that "[i]n another embodiment, minions 124 use accounts at different Internet Service Providers." EX.1004, 16:27-28; EX.1003, ¶¶155-156.

d. [Claim 16g]

Petitioner incorporates by reference its discussion in Section VI.A.2.g herein. Thus, for the reasons discussed above, claim 16 is obvious under Cousins in view of the common knowledge of a POSA. EX.1003, ¶171.

3. Claim 19

a. [Claim 19a], [Claim 19b], [Claim 19c], [Claim 19d], and [Claim 19e]

Petitioner incorporates by reference its discussion in Section VI.A.3.a, VI.A.3.b, VI.A.3.c, VI.A.3.d, and VI.A.3.e herein.

b. [Claim 19f]

Petitioner incorporates by reference its discussion in Section VI.A.1.f herein.

If the Board does not agree that Cousins discloses "a proxy" then such a limitation would be common knowledge to a POSA. EX.1003, ¶177.

Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2. In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶177. Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, ¶177. In

addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶177. As would have been known by a POSA, Squid is "a popular open source caching proxy for the Web." EX.1009, 1; EX.1003, ¶178.

It would have been obvious to a POSA to utilize minions as proxies or to have minions use proxies as Cousins specifically states such a configuration. EX.1003, ¶179. A POSA would also have known of the advantages of using proxies, for example as stated in applicant's own admitted prior art in the '770 patent, "[p]roxies can be used to bypass certain Internet restrictions (*e.g*. firewalls) by enabling a user to request the content through a (remote) proxy server instead of accessing the content directly" where "[p]roxy servers are often used to get around geo-IP based content restrictions." EX.1004, 2:25-29; EX.1003, ¶179.

Dr. Williams concludes it would have been obvious to adapt Cousins with the common knowledge of a POSA. EX.1003, ¶180.

Thus, for the reasons discussed above, claim 19 is obvious under Cousins in view of the common knowledge of a POSA. EX.1003, ¶181.

4. Claim 2

> *"… further comprising: checking and analyzing the metadata of the reply received from the target."*

Petitioner incorporates by reference its discussion in Section VI.B.1 herein.

If the Board does not agree that Cousins discloses "checking and analyzing the metadata" then such a limitation would be common knowledge to a POSA. EX.1003, ¶183.

Cousins discloses that "[c]rawlers gather this information by using the Internet in a process known as crawling" where "[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet." EX.1004, 1:17-20; EX.1003, ¶184.

The '770 patent states that "[i]t should be noted that the metadata of the response is always visible to the Request Optimizer 104, e.g. URL and HTTP headers." EX.1001, 7:18-20; EX.1003, ¶185.

A POSA would have understood that the use of standard web protocols includes checking and analyzing associated metadata found in URL and HTTP headers. EX.1003, ¶173. For example, a POSA would have known that, as stated in RFC7231, that when a network "message includes a payload body, the representation header fields describe how to interpret the representation data enclosed in the payload body." EX.1005, 5; EX.1003, ¶186.

A POSA would have known that, as stated in RFC7231, "HTTP semantics include the intentions defined by each request method … and the meaning of other control data and resource metadata that might be given in response header fields." EX.1005, 6; EX.1003, ¶187. Dr. Williams concludes it would have been obvious to adapt Cousins with the common knowledge of a POSA. EX.1003, ¶¶188-189.

5. Claim 3

> *"… further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata."*

Petitioner incorporates by reference its discussion in Sections VI.B.1 and VI.B.4 herein.

If the Board does not agree that Cousins discloses "forwarding the reply to the user device pursuant to the analysis of the metadata" then such a limitation would be common knowledge to a POSA. EX.1003, ¶190.

As discussed, Cousins discloses "[a] request source [that] is any computer process that requests information from a data address" where "a user will access the crawler's web site through the user's Internet connection, and then input a request for a specific item into a form provided by the crawler's be site" EX.1004, 6:15-24; EX.1003, ¶191. Cousins discloses "scheduled" and "user-demanded crawl[s]." EX.1004, 3:26; EX.1003, ¶191. As would have been obvious to a POSA, Cousins disclosure of a user's access through an Internet connection, and to input a request,

as well as the creation of a user-demanded crawl discloses requests and replies to and from a user device. EX.1003, ¶191.

As discussed in claim 2, Cousins discloses that "[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet." EX.1004, 1:17-20; EX.1003, ¶192. And, that the '770 patent states that "[i]t should be noted that the metadata of the response is always visible to the Request Optimizer 104, e.g. URL and HTTP headers." EX.1001, 7:18-20; EX.1003, ¶192.

A POSA would have understood that the use of standard web protocols includes forwarding content such as the reply to the user device based on the analysis of the metadata. EX.1003, ¶193. For example, a POSA would have known that, as stated in RFC7231, the HTTP protocol allows for "the recipient [to] establish a tunnel to the destination origin server identified by the request-target and, if successful, thereafter restrict its behavior to blind forwarding of packets, in both directions, until the tunnel is closed." EX.1005, 30; EX.1003, ¶193. Dr. Williams concludes it would have been obvious to adapt Cousins with the common knowledge of a POSA. EX.1003, ¶¶194-195.

6. Claim 17

Petitioner incorporates by reference its discussion in Sections VI.A.6, VI.A.13, and VI.B.2 herein. *See* EX.1003, ¶196.

7. Claim 18

> *"… wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device."*

Petitioner incorporates by reference its discussion in Section VI.B.2 herein.

If the Board does not agree that Cousins discloses "wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device" then such a limitation would be common knowledge to a POSA. EX.1003, ¶197.

A POSA would have understood that the use of standard web protocols includes utilizing the same proxy for an unlimited number of requests received from the same user device, also known as a persistent connection. For example, a POSA would have known that, as stated in RFC7540, the HTTP specifies that "HTTP/2 connections are persistent" and that "[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or until the server closes the connection." EX.1008, 65;

EX.1003, ¶198. Dr. Williams concludes it would have been obvious to adapt Cousins with the common knowledge of a POSA. EX.1003, ¶¶199-200.

8. Claim 20

> *"… further comprising checking and analyzing the metadata of the reply received from the target."*

Petitioner incorporates by reference its discussion in Sections VI.B.3 and VI.B.4 herein.

If the Board does not agree that Cousins discloses "further comprising checking and analyzing the metadata of the reply received from the target" then such a limitation would be common knowledge to a POSA as discussed above regarding claim 2. *See* Section VI.B.4; EX.1003, ¶¶201-202.

9. Claim 21

> *"… further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata."*

Petitioner incorporates by reference its discussion in Sections VI.B.8, VI.B.3, and VI.B.5 herein.

If the Board does not agree that Cousins discloses "further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata" then

such a limitation would be common knowledge to a POSA as discussed above regarding claim 3. *See* Section VI.B.5; EX.1003, ¶¶203-204.

C. Ground 3: Obviousness of Claims 1, 7, 8, 13, and 16-19 by Cousins in View of Taylor

1. The Combination of Cousins and Taylor

Cousins discloses the use of minions to retrieve crawl requests and crawl the target web sites. *See* EX.1004, Abstract, code (57). Cousins discloses that the minions "use accounts at different Internet Service Providers" and that minions "may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies." EX.1004, 16:27-29; EX.1003, ¶206.

A POSA would have recognized that there are advantages to using proxies from Internet Service Providers that allows for a large network to more effectively crawl and retrieve information from target web sites. EX.1003, ¶207. Thus, it would have been obvious to adapt and expand Cousins to incorporate the service provider proxies and infrastructure taught by Taylor. *Id*.

Taylor discloses a "service provider device [that] is further configured to direct network communication to a proxy device which, in turn, exchanges communication with targeted server devices over a wide area network, such as the Internet." EX.1006, ¶[0008]; EX.1003, ¶208. Taylor discloses that the service provider device is configured to "detect the communication type identifier"

associated with the request and based on a correspondence of the destination IP address and a DNS database the service provider device "can select an appropriate middlebox 40" to proxy the network communication between a communication device and a server device. EX.1006, ¶[0044]; EX.1003, ¶208.

The use of a service provider proxies and databases as suggested by Taylor into the Cousins' web crawler architecture would have been an obvious implementation for a POSA and would have also been done expecting success. EX.1003, ¶209.

Cousins and Taylor provide a teaching, suggestion, or motivation to combine Taylor's service provider middlebox proxy architecture with Cousins web crawler architecture that also uses a database to select the appropriate proxies associated with a target server. EX.1003, ¶210.

Such a combination represents combining prior art elements (service provider proxies) according to known methods (web page crawler requests and responses) to yield predictable results (increased amounts of proxies and performance). EX.1003, ¶211.

Cousins in view of Taylor discloses each of the limitations of claims 1, 7, 8, 13, and 16-19 of the '770 Patent.

2. Claim 1

See EX.1003, ¶¶214-229.

a. [Claim 1a], [Claim 1b], [Claim 1c], [Claim 1d], and [Claim 1e]

Petitioner incorporates by reference its discussion in Sections VI.A.1.a, VI.A.1.b, VI.A.1.c, VI.A.1.d, VI.A.1.e and VI.B.1.a, VI.B.1.b, VI.B.1.c, VI.B.1.d, VI.B.1.e herein.

b. [Claim 1f]

Petitioner incorporates by reference its discussion in Sections VI.A.1.f and VI.B.1.f herein.

If the Board does not agree that Cousins discloses "a proxy" then such a limitation would be obvious under Cousins in view of Taylor. EX.1003, ¶222.

Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, ¶223. In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." *Id,* 16:28-29. Squid, as known by a POSA, is a proxy rotator. EX.1009, 1-2; EX.1003, ¶223.

Taylor discloses selecting a proxy based on the strategy of the type of communication detected, for example, "[b]ased upon the detected communication

type identifier 32, the service provider device 12 can select an appropriate, associated middlebox 40 configured to proxy and monitor the application, workflow, or protocol traffic from the client device 15." EX.1006, ¶[0029] ; EX.1003, ¶224.

In another example, Taylor discloses selecting a proxy based upon the type of application, workflow, or protocol being used, for example, each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis." EX.1006, ¶[0024]; EX.1003, ¶225. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶226-227.

c. [Claim 1g]

Petitioner incorporates by reference its discussion in Sections VI.A.1.g and VI.B.1.g herein.

Thus, for the reasons discussed above, claim 1 is obvious under Cousins in view of Taylor. EX.1003, ¶229.

3. Claim 16

See EX.1003, ¶¶230-244.

a. [Claim 16a] , [Claim 16b], [Claim 16d], [Claim 16e], and [Claim 16f]

Petitioner incorporates by reference its discussion in Sections VI.A.2.a, VI.A.2.d, VI.A.2.e, and VI.A.2.f and VI.B.2.a, VI.A.2.b, VI.B.2.d, VI.B.2.e, and VI.B.2.f herein.

b. [Claim 16c]

Petitioner incorporates by reference its discussion in Sections VI.A.2.c and VI.B.2.c herein.

If the Board does not agree that Cousins discloses a request optimizer that is a part of the service provider infrastructure then Taylor, in the same field, discloses such a limitation. EX.1003, ¶233.

Taylor discloses a "proxy that is a part of the service provider infrastructure that accepts requests from the User Device and implements the requests through the system." EX.1003, ¶234. For example, Taylor discloses "a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel" including a "virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox." EX.1006, ¶[0007]; EX.1003, ¶234. Taylor discloses that the "service provider device is further configured to direct network communication to a proxy device which, in turn,

exchanges communication with targeted server devices over a wide area network, such as the Internet." EX.1006, ¶[0008]; EX.1003, ¶234. Taylor discloses a method that "includes receiving, by the service provider device, a device access request from the client device … [and] establishing, by the service provider device, a secure communication channel between the client device and a middlebox associated with the detected communication type identifier. EX.1006, ¶[0009]; EX.1003, ¶234. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶235-236.

c. [Claim 16g]

Petitioner incorporates by reference its discussion in Sections VI.A.2.g and VI.B.2.g herein.

If the Board does not agree that Cousins, or Cousins in view of the common knowledge of a POSA, discloses "a proxy selected from a proxy database" then such a limitation would be obvious under Cousins in view of Taylor as discussed above regarding claim 7. EX.1003, ¶241. Taylor discloses that the service provider device, *e.g.*, the service provider infrastructure, is configured to "detect the communication type identifier" associated with the request and based on a correspondence of the destination IP address and a DNS database the service provider device "can select an appropriate middlebox 40" to proxy the network communication between a communication device and a server device. EX.1006, ¶[0044]; EX.1003, ¶241. Thus,

Taylor discloses the use of a database to select a proxy, *e.g.*, middlebox 40 according to the scraping strategy. EX.1003, ¶241. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶242-243.

Thus, for the reasons discussed above, claim 16 is obvious under Cousins in view of Taylor. EX.1003, ¶244.

4. Claim 19

See EX.1003, ¶¶245-255.

a. [Claim 19a] , [Claim 19b], [Claim 19c], [Claim 19d], and [Claim 19e]

Petitioner incorporates by reference its discussion in Sections VI.A.3.a, VI.A.3.b, VI.A.3.c, VI.A.3.d, and VI.A.3.e and VI.B.3.a, VI.B.3.b, VI.B.3.c, VI.B.3.d, and VI.B.3.e herein.

b. [Claim 19f]

Petitioner incorporates by reference its discussion in Sections VI.A.3.f and VI.B.3.f herein.

If the Board does not agree that Cousins discloses "a proxy" then such a limitation would be obvious under Cousins in view of Taylor. EX.1003, ¶250.

Cousins discloses that "in general, minions 124 are located topologically as close as possible to the target site's server 708." EX.1004, 15:1-2; EX.1003, ¶225.

In addition, "[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶225. Squid, as known by a POSA, is a proxy rotator. EX.1009, 1-2; EX.1003, ¶251.

Taylor discloses selecting a proxy based on the strategy of the type of communication detected, for example, "[b]ased upon the detected communication type identifier 32, the service provider device 12 can select an appropriate, associated middlebox 40 configured to proxy and monitor the application, workflow, or protocol traffic from the client device 15." EX.1006, ¶[0029]; EX.1003, ¶252.

In another example, Taylor discloses selecting a proxy based upon the type of application, workflow, or protocol being used, for example, "each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis." EX.1006, ¶[0024] ; EX.1003, ¶253. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶254.

Thus, for the reasons discussed above, claim 19 is obvious under Cousins in view of Taylor. EX.1003, ¶255.

5. Claim 7

> *"… wherein the proxy is selected by a service provider infrastructure from a proxy database."*

Petitioner incorporates by reference its discussion in Section VI.C.2 herein.

If the Board does not agree that Cousins discloses a proxy "selected by a service provider infrastructure from a proxy database" then Taylor, in the same field, discloses such a limitation. EX.1003, ¶257. Taylor discloses "a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel" including a "virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox." EX.1006, ¶[0007]; EX.1003, ¶258. Taylor discloses that the "service provider device is further configured to direct network communication to a proxy device which, in turn, exchanges communication with targeted server devices over a wide area network, such as the Internet." EX.1006, ¶[0008]; EX.1003, ¶258. Taylor discloses a method that "includes receiving, by the service provider device, a device access request from the client device … [and] establishing, by the service provider device, a secure communication channel between the client device and a middlebox associated with the detected communication type identifier. EX.1006, ¶[0009]; EX.1003, ¶258.

Taylor discloses that the service provider device, *e.g.*, the service provider infrastructure, is configured to "detect the communication type identifier" associated with the request and based on a correspondence of the destination IP address and a DNS database the service provider device "can select an appropriate middlebox 40" to proxy the network communication between a communication device and a server device. EX.1006, ¶[0044]; EX.1003, ¶259. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶260-261.

6. Claim 8

> *"… wherein the proxy is selected according to the scraping strategy and includes geolocation settings, proxy type, proxy availability, or any combination thereof."*

Petitioner incorporates by reference its discussion in Sections VI.C.2 and VI.C.5 herein.

Cousins discloses that policies may include geolocation settings, for example, "[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc." EX.1004, 3:15-16; EX.1003, ¶262. Furthermore, Cousins discloses that "certain types of requests may be transmitted to a specific minion 128" where " minion 124 may be restricted to a particular area (such as 'stamps & coins') so that the behavior of the minion 124 will emulate human behavior." EX.1004, 16:18-21; EX.1003, ¶234. Cousins discloses the minions "may

also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid)." EX.1004, 16:28-29; EX.1003, ¶262.

Taylor discloses selecting proxies, e.g., middleboxes, based on proxy type. For example, "the middlebox is configured to monitor and control the network traffic associated with a particular application, workflow, or protocol as executed by the client device." EX.1006, ¶[0017]; EX.1003, ¶263. Taylor discloses selecting the proxy based on the type of application. For example, "each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis." In addition, "the first middlebox 40-1 can be configured to proxy and monitor network traffic pertaining to a particular type of video conferencing (e.g., SKYPE) while the second middlebox 40-2 can be configured to proxy and monitor network traffic pertaining to a particular type motion picture streaming (e.g., NETFLIX)." EX.1006, ¶[0024]; EX.1003, ¶263.

Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation. EX.1003, ¶264.

7. Claim 13

Petitioner incorporates by reference its discussion in Section VI.C.2 herein.

238. Taylor discloses "a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel" including a "virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox." EX.1006, ¶[0007]; EX.1003, ¶266. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶267-268.

8. Claim 15

> *"… wherein the proxy is a proxy server integrated into a service provider infrastructure."*

Petitioner incorporates by reference its discussion in Section VI.C.2 herein.

Taylor discloses "a service provider device" as a part of the service provider infrastructure that performs an adjusting of a parameter autonomously from the user device. EX.1003, ¶270. For example, a user device request to a server device may be redirected to a server provider device. For example, " assume the case where a user of a first computer device 18-1 seeks to establish a video conferencing connection (e.g., SKYPE) with server device 22." "In such a case, the computer device 18-1 generates a device access request 30 identifying the server device 22 and forwards the request within the residential network 14 to the data communication device 16," but "based upon the configuration of the data

communication device 16 (e.g., as configured with the OpenFlow firmware), the data communication device 16 forwards the device access request 30 to the service provider device 12 via network communication channel 25 (e.g., over a WAN such as the Internet), rather than to the server device 22 identified by the device access request 30." EX.1006, [0027]; EX.1003, ¶270. Dr. Williams concludes it would have been obvious to adapt Cousins with the teachings of Taylor. EX.1003, ¶¶271-272.

9. Claim 17

Petitioner incorporates by reference its discussion in Sections VI.A.6, VI.B.6, and VI.C.3 herein.

10. Claim 18

Petitioner incorporates by reference its discussion in Sections VI.B.7 and VI.C.3 herein.

D. Ground 4: Obviousness of Claims 2, 3, 20, and 21 Based on Cousins in View of RFC7231

1. The Combination of Cousins and RFC7231

Cousins discloses the use of minions to retrieve crawl requests and crawl the target web sites. *See* EX.1004, Abstract, code (57); EX.1003, ¶278. Cousins discloses that "[c]rawlers gather this information by using the Internet in a process known as crawling" and that "[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols

(TCP/IP, HTTP, etc.) over the internet." EX.1004, 1:17-20; EX.1003, ¶278. RFC7231 is the standards track document to define the Hypertext Transfer Protocol (HTTP/1.1) semantics and content and thus are highly relevant to the protocol features disclosed in Cousins. EX.1003, ¶278. Thus, it would have been obvious to combine Cousins and RFC7231. *Id,* ¶279.

Cousins is directed to accessing information from target web sites using standard web protocols, including TCP/IP and HTTP over the Internet. *Id,* ¶280. A POSA would have recognized the use of RFC7231 standards in the use of Internet communications. *Id.* RFC7231 specifies standards track protocols for the Internet community as it describes the usage of metadata, request methods, header fields, responses, and security considerations. EX.1005, 3-5; EX.1003, ¶280.

Cousins and RFC7231 provide a teaching, suggestion, or motivation to combine Cousins with RFC7231 standards for consistency and interoperability with other network devices and applications. *Id,* ¶281.

Such a combination represents combining prior art elements (network components including clients, proxies, and target servers) according to known methods (Internet standards) to yield predictable results (compatibility and interoperability). *Id,* ¶282. The combination is merely a simple modification of one known element (network communications) for another (standards-based

communication). *Id.* The combination is the use of a known technique (network communications) to control similar devices (standards) in the same way. *Id.*

Cousins in view of RFC7231 discloses each of the limitations of claims 2, 3, 20, and 21 of the '770 Patent as discussed below. *Id*, ¶¶283-284. These claims are obvious by the disclosures in Cousins and RFC7231.

2. Claims 2 and 20

> *"... further comprising: checking and analyzing the metadata of the reply received from the target."*

Petitioner incorporates by reference its discussion in Sections VI.A.1, VI.A.3, VI.B.1, VI.B.3, VI.C.2, and VI.C.4 herein.

Cousins discloses that "computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the 20 internet." EX.1004, 1:17-20; EX.1003, ¶285.

RFC7231 discloses "[f]or the purposes of HTTP, a 'representation' is information that is intended to reflect a past, current, or desired state of a given resource, in a format that can be readily communicated via the protocol, and that consists of a set of representation metadata and a potentially unbounded stream of representation data." EX.1005, 7; EX.1003, ¶286. RFC7231 explains that the metadata is analyzed to determine how to interpret data, for example,

"[r]epresentation header fields provide metadata about the representation" and "[w]hen a message includes a payload body, the representation header fields describe how to interpret the representation data enclosed in the payload body." EX.1005, 8; EX.1003, ¶286.

Thus, claims 2 and 20 are obvious under Cousins in view of RFC7231. EX.1003, ¶¶287-288.

3. Claims 3 and 21

> *"… further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata."*

Petitioner incorporates by reference its discussion in Section VI.D.2 herein.

Cousins discloses that "computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the 20 internet." EX.1004, 1:17-20; EX.1003, ¶289.

RFC7231 discloses analysis of metadata that includes where "HTTP semantics include the intentions defined by each request method (Section 4), extensions to those semantics that might be described in request header fields … and the meaning of other control data and resource metadata that might be given in response header fields." EX.1005, 6; EX.1003, ¶290. RFC7231 describes a tunnel mode as bi-directional where "[t]he CONNECT method requests that the recipient

establish a tunnel to the destination origin server identified by the request-target and, if successful, thereafter restrict its behavior to blind forwarding of packets, in both directions, until the tunnel is closed." EX.1005, 30; EX.1003, ¶290.

RFC7231 explains forwarding the response by establishing the tunnel based on the metadata header analysis, for example, where "[t]he recipient proxy can establish a tunnel either by directly connecting to the request-target or, if configured to use another proxy, by forwarding the CONNECT request to the next inbound proxy" where a successful "response indicates that the sender (and all inbound proxies) will switch to tunnel mode immediately after the blank line that concludes the successful response's header section." EX.1005, 31; EX.1003, ¶291.

Thus, claims 3 and 21 are obvious under Cousins in view of RFC7231. EX.1003, ¶¶292-293.

E. Ground 5: Obviousness of Claim 18 Based on Cousins in View of RFC7540

1. The Combination of Cousins and RFC7540

Cousins has been discussed above. *See* Section VI.D.1; EX.1004, 1:17-20; EX.1003, ¶295. RFC7540 is the standards track document to define the Hypertext Transfer Protocol (HTTP/2) semantics and thus are highly relevant to the protocol features disclosed in Cousins. EX.1003, ¶295. Thus, it would have been obvious to combine Cousins and RFC7540. EX.1003, ¶296.

Cousins is directed to accessing information from target web sites using standard web protocols, including TCP/IP and HTTP over the Internet. EX.1003, ¶297. A POSA would have recognized the use of RFC7540 standards in the use of Internet communications. *Id.* RFC7540 specifies standards track protocols for the Internet community as it describes the usage of protocol, frames, streams and multiplexing, message exchanges, connection management, and security considerations. EX.1008, 2-4; EX.1003, ¶297. Thus, it would have been obvious to adapt Cousins to adhere to the RFC7540 standards that entails network communications. EX.1003, ¶297.

Cousins and RFC7540 provide a teaching, suggestion, or motivation to combine Cousins with RFC7540 standards for consistency and interoperability with other network devices and applications. *Id*, ¶298.

Such a combination represents combining prior art elements (network components including clients, proxies, and target servers) according to known methods (Internet standards) to yield predictable results (compatibility and interoperability). *Id*, ¶299. The combination is merely a simple modification of one known element (network communications) for another (standards-based communication). *Id.* The combination is the use of a known technique (network communications) to control similar devices (standards) in the same way. *Id.* The

limitations of claim 18 are obvious to a POSA based on the disclosures in Cousins and RFC7540. *Id*, ¶300.

1. Claim 18

> *"… wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device."*

Petitioner incorporates by reference its discussion in Sections VI.A.2, VI.B.2, and VI.C.3 herein.

Cousins discloses that "computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the 20 internet." EX.1004, 1:17-20; EX.1003, ¶301.

RFC7540 discloses connection management that includes persistent connections where "HTTP/2 connections are persistent" and that "[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or HTTP/2 connections are persistent" and that "[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or until the server closes the connection." EX.1008, 65; EX.1003, ¶302.

Thus, claim 18 is obvious under Cousins in view of RFC7231. EX.1003, ¶¶303-304.

VII. CONCLUSION

Petitioner respectfully requests cancellation of the Challenged Claims of the ’770 Patent.

Respectfully submitted,

Date: May 11, 2023

By: /s/ *Thomas M. Dunham*
Thomas M. Dunham
Reg. No. 39,965

Cherian LLP
1901 L Street NW, Suite 700
Washington, D.C. 20036
(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.

CERTIFICATE OF COMPLIANCE WITH TYPE-VOLUME LIMITS

This petition consists of 13,993 words, which is no more than 14,000 words, excluding table of contents, table of authorities, table of exhibits, mandatory notices, certificate of service, and this certificate. The petition complies with the type-volume as mandated in 37 C.F.R. §42.24. In preparing this certificate, counsel has relied on the word count of the word-processing system used to prepare the paper (Microsoft Word).

Respectfully submitted,

Date: May 11, 2023

By: /s/ *Thomas M. Dunham*
Thomas M. Dunham
Reg. No. 39,965

Cherian LLP
1901 L Street NW, Suite 700
Washington, D.C. 20036
(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §42.6(e) and 37 C.F.R. §42.105(a), the undersigned hereby certifies the PETITION FOR *INTER PARTES* REVIEW and all exhibits thereto were served on the undersigned date via FedEx Express, at the following addresses:

Patent File Correspondence Address:
Oxylabs, UAB
ATTN: Julius Cerniauskas
Svitrigailos G. 32
Vilnius, LT-03230
LITHUANIA

IPR Counsel Address:
Eagle Robinson
Norton Rose Fulbright US LLP
98 San Jacinto Boulevard, Suite 110
Austin, TX 78701

Daniel Leventhal
Norton Rose Fulbright US LLP
Fulbright Tower
1301 McKinney Street, Suite 5100
Houston, TX 77010

Courtesy copies were also served via e-mail at the following addresses:

Eagle.robinson@nortonrosefulbright.com

Daniel.leventhal@nortonrosefulbright.com

Respectfully submitted,

Date: May 11, 2023

By: /s/ *Thomas M. Dunham*
Thomas M. Dunham
Reg. No. 39,965

Cherian LLP
1901 L Street NW, Suite 700
Washington, D.C. 20036
(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.