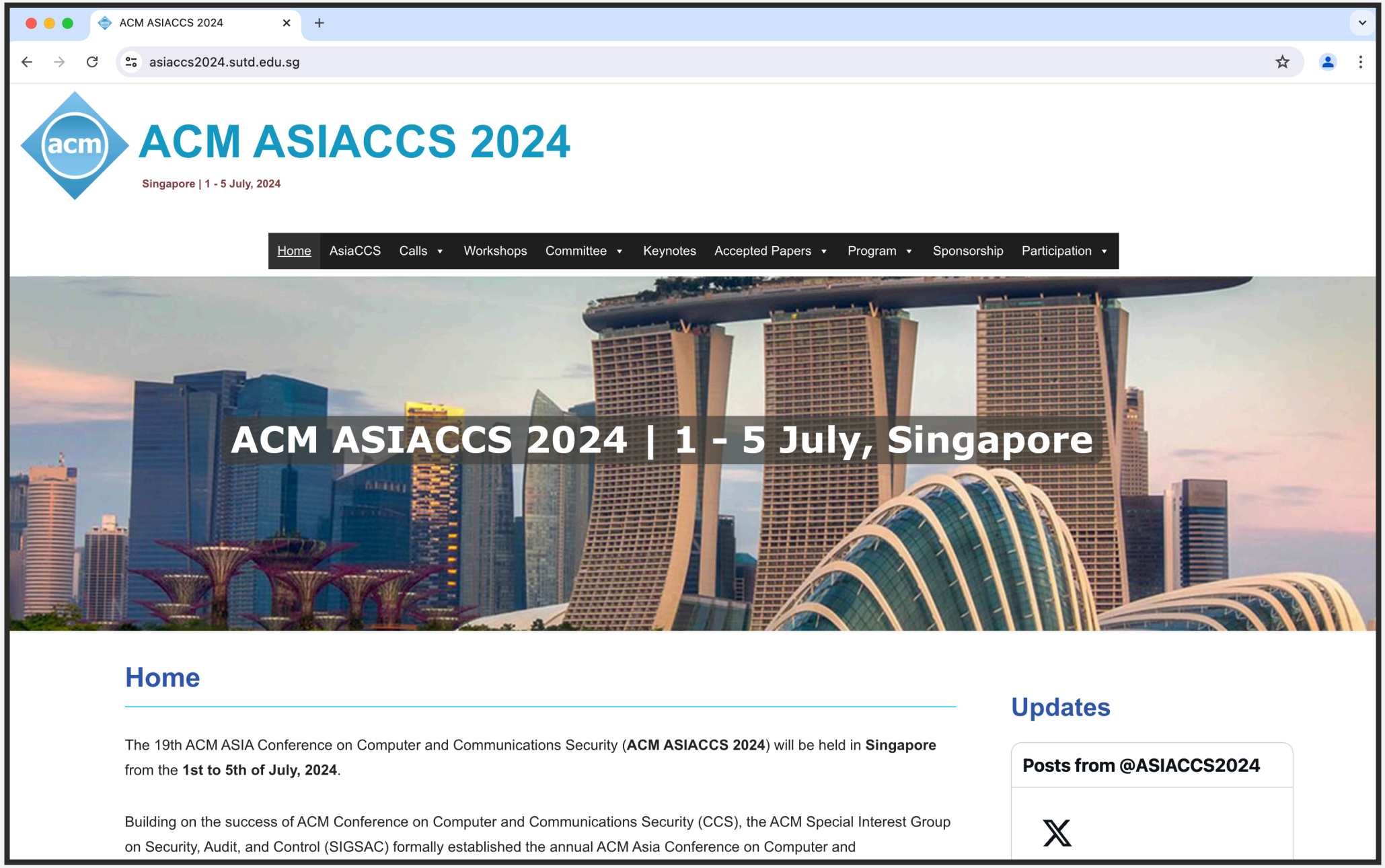
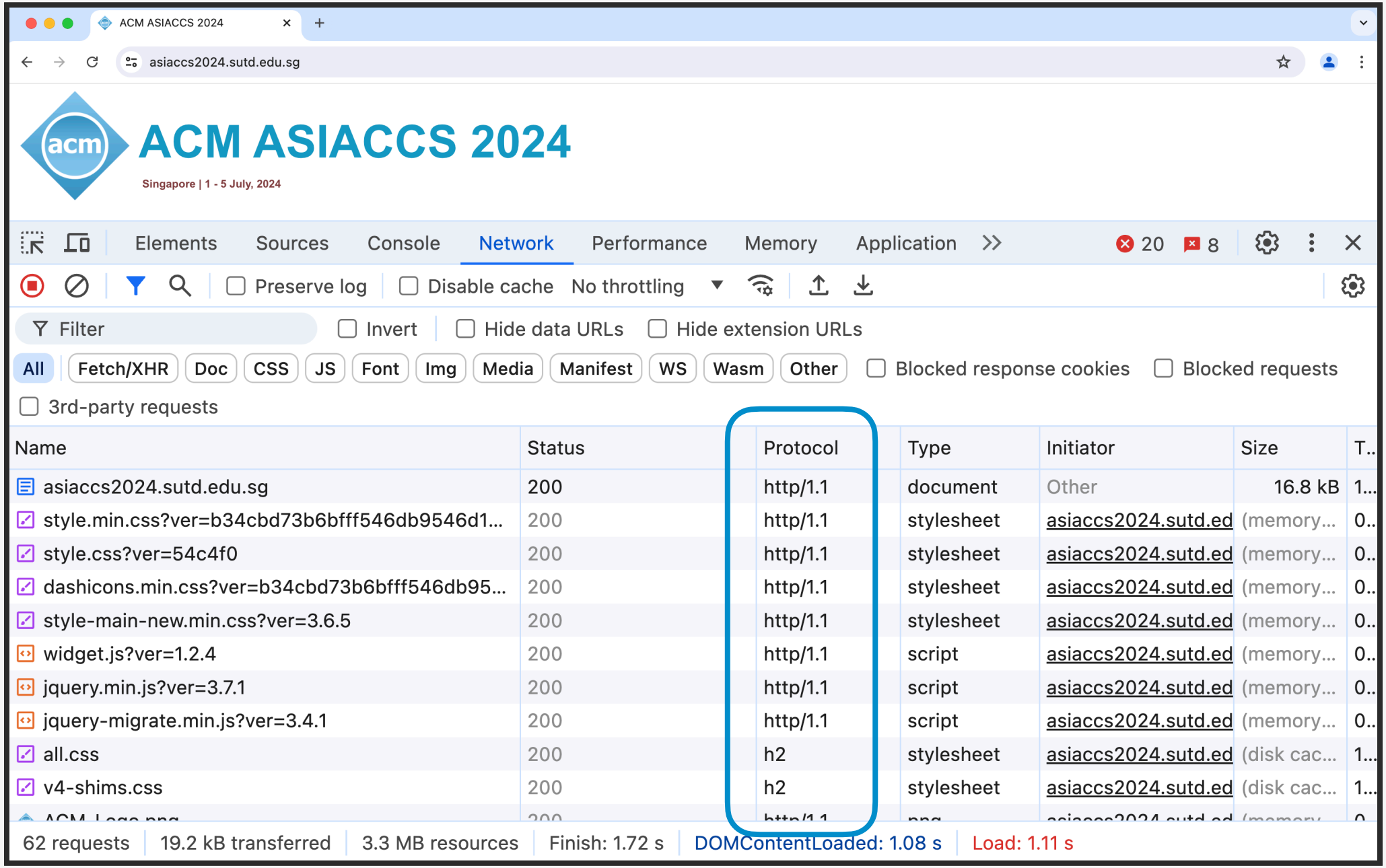
CISPA
HELMHOLTZ CENTER FOR INFORMATION SECURITY

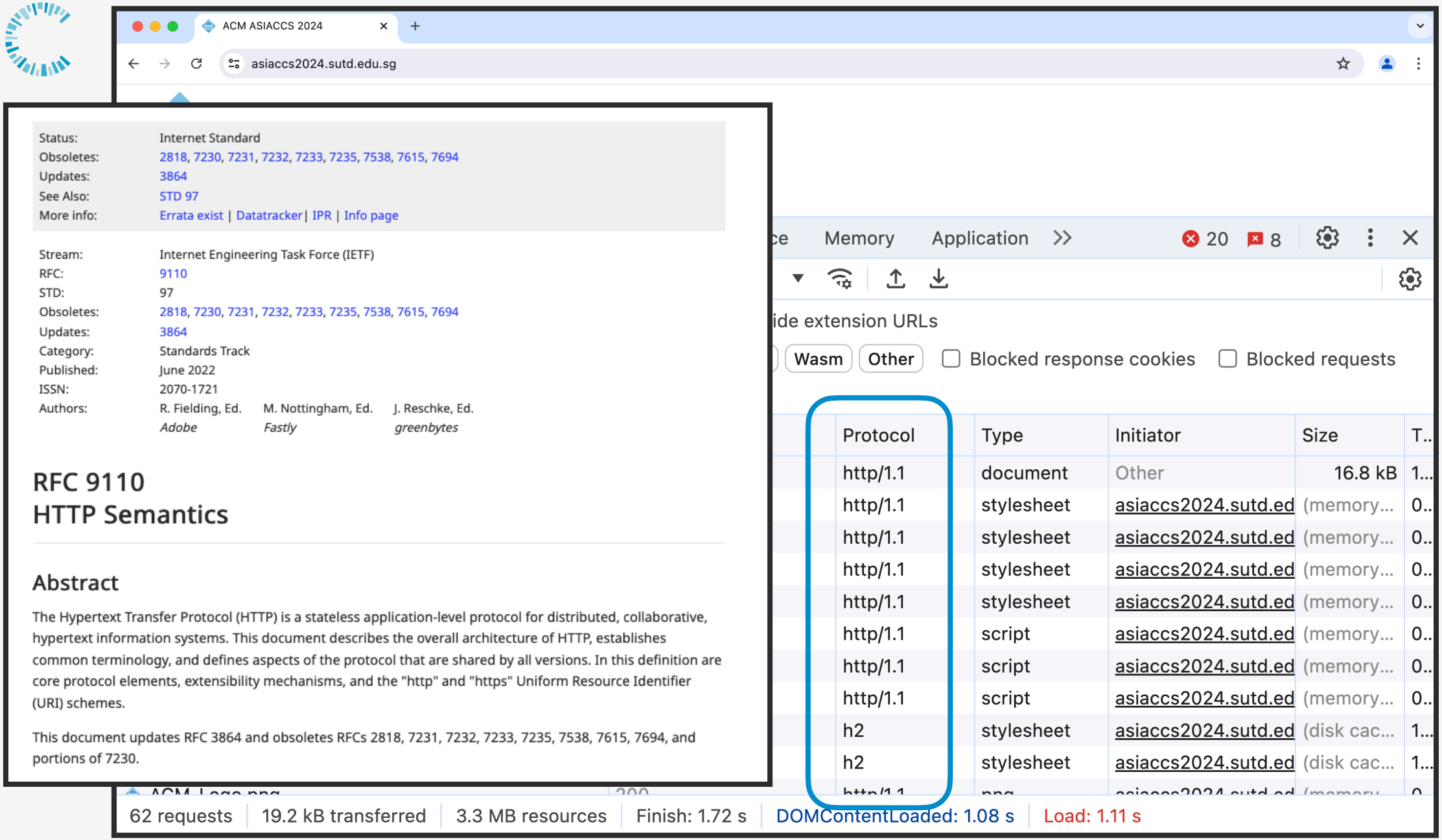
Who's Breaking the Rules? Studying Conformance to the HTTP Specifications and its Security Impact

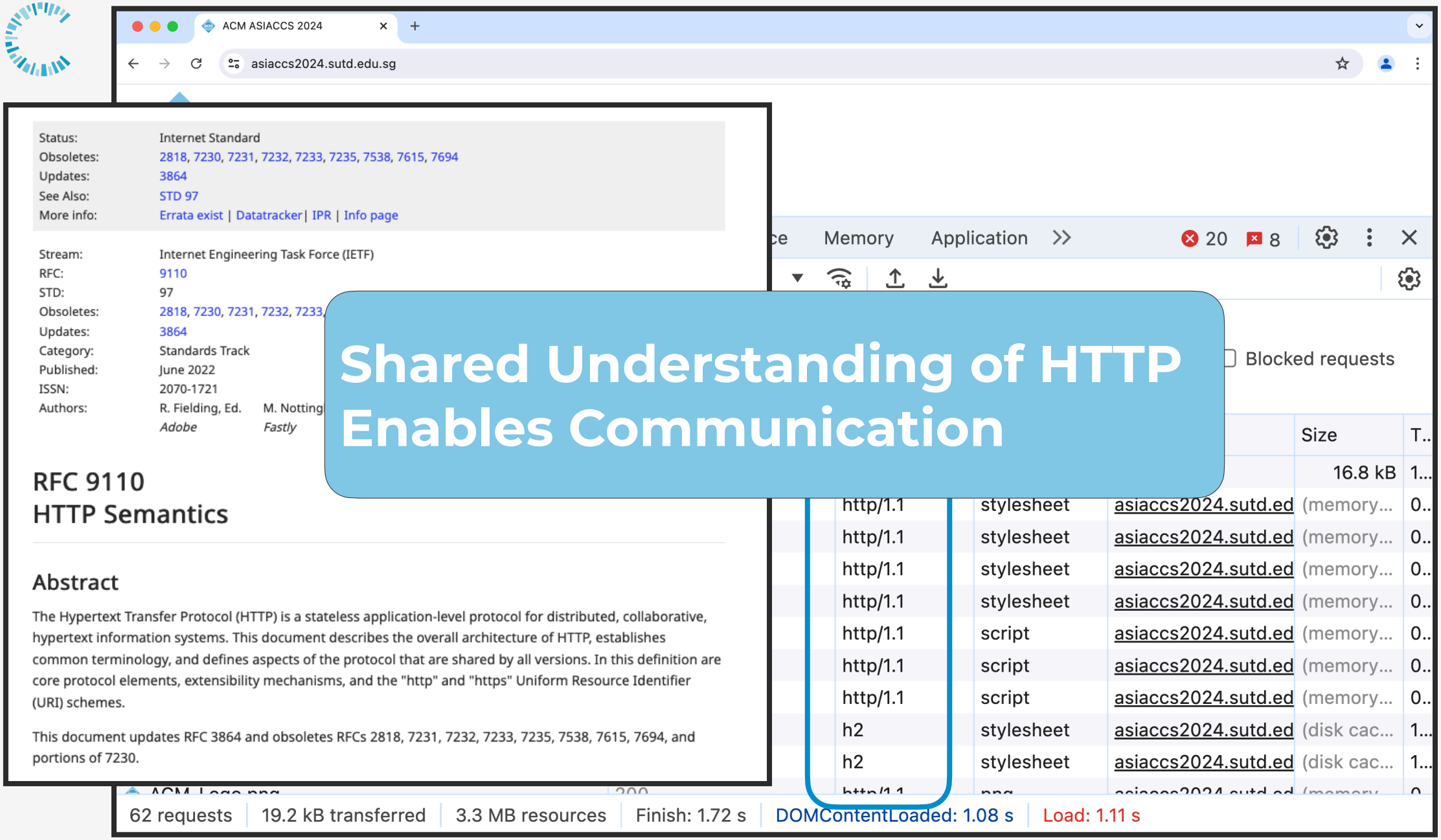
Jannis Rautenstrauch, *Ben Stock*
CISPA Helmholtz Center for Information Security
jannis.rautenstrauch@cispa.de

19th ACM ASIA Conference on Computer and Communications Security 2024


ACM ASIACCS 2024
asiaccs2024.sutd.edu.sg
ACM ASIACCS 2024
Singapore | 1 - 5 July, 2024
Home
AsiaCCS
Calls
Workshops
Committee
Keynotes
Accepted Papers
Program
Sponsorship
Participation
ACM ASIACCS 2024 | 1 - 5 July, Singapore
Home
The 19th ACM ASIA Conference on Computer and Communications Security (ACM ASIACCS 2024) will be held in Singapore from the 1st to 5th of July, 2024.
Building on the success of ACM Conference on Computer and Communications Security (CCS), the ACM Special Interest Group on Security, Audit, and Control (SIGSAC) formally established the annual ACM Asia Conference on Computer and
Updates
Posts from @ASIACCS2024


ACM ASIACCS 2024
asiaccs2024.sutd.edu.sg
ACM ASIACCS 2024
Singapore | 1 - 5 July, 2024
Elements
Sources
Console
Network
Performance
Memory
Application
20
8
Preserve log
Disable cache
No throttling
Filter
Invert
Hide data URLs
Hide extension URLs
All
Fetch/XHR
Doc
CSS
JS
Font
Img
Media
Manifest
WS
Wasm
Other
Blocked response cookies
Blocked requests
3rd-party requests
Name
Status
Protocol
Type
Initiator
Size
asiaccs2024.sutd.edu.sg 200 http/1.1 document Other 16.8 kB
style.min.css?ver=b34cbd73b6bfff546db9546d1... 200 http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
style.css?ver=54c4f0 200 http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
dashicons.min.css?ver=b34cbd73b6bfff546db95... 200 http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
style-main-new.min.css?ver=3.6.5 200 http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
widget.js?ver=1.2.4 200 http/1.1 script asiaccs2024.sutd.ed (memory...
jquery.min.js?ver=3.7.1 200 http/1.1 script asiaccs2024.sutd.ed (memory...
jquery-migrate.min.js?ver=3.4.1 200 http/1.1 script asiaccs2024.sutd.ed (memory...
all.css 200 h2 stylesheet asiaccs2024.sutd.ed (disk cac...
v4-shims.css 200 h2 stylesheet asiaccs2024.sutd.ed (disk cac...
62 requests
19.2 kB transferred
3.3 MB resources
Finish: 1.72 s
DOMContentLoaded: 1.08 s
Load: 1.11 s


ACM ASIACCS 2024
asiaccs2024.sutd.edu.sg
Status: Internet Standard
Obsoletes: 2818, 7230, 7231, 7232, 7233, 7235, 7538, 7615, 7694
Updates: 3864
See Also: STD 97
More info: Errata exist | Datatracker | IPR | Info page
Stream: Internet Engineering Task Force (IETF)
RFC: 9110
STD: 97
Obsoletes: 2818, 7230, 7231, 7232, 7233, 7235, 7538, 7615, 7694
Updates: 3864
Category: Standards Track
Published: June 2022
ISSN: 2070-1721
Authors: R. Fielding, Ed. Adobe; M. Nottingham, Ed. Fastly; J. Reschke, Ed. greenbytes
RFC 9110
HTTP Semantics
Abstract
The Hypertext Transfer Protocol (HTTP) is a stateless application-level protocol for distributed, collaborative, hypertext information systems. This document describes the overall architecture of HTTP, establishes common terminology, and defines aspects of the protocol that are shared by all versions. In this definition are core protocol elements, extensibility mechanisms, and the "http" and "https" Uniform Resource Identifier (URI) schemes.
This document updates RFC 3864 and obsoletes RFCs 2818, 7231, 7232, 7233, 7235, 7538, 7615, 7694, and portions of 7230.
ce
Memory
Application
20
8
ide extension URLs
Wasm
Other
Blocked response cookies
Blocked requests
Protocol
Type
Initiator
Size
http/1.1 document Other 16.8 kB
http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
http/1.1 stylesheet asiaccs2024.sutd.ed (memory...
http/1.1 script asiaccs2024.sutd.ed (memory...
http/1.1 script asiaccs2024.sutd.ed (memory...
http/1.1 script asiaccs2024.sutd.ed (memory...
h2 stylesheet asiaccs2024.sutd.ed (disk cac...
h2 stylesheet asiaccs2024.sutd.ed (disk cac...
62 requests
19.2 kB transferred
3.3 MB resources
Finish: 1.72 s
DOMContentLoaded: 1.08 s
Load: 1.11 s


ACM ASIACCS 2024
asiaccs2024.sutd.edu.sg
Status: Internet Standard
Obsoletes: 2818, 7230, 7231, 7232, 7233, 7235, 7538, 7615, 7694
Updates: 3864
See Also: STD 97
More info: Errata exist | Datatracker | IPR | Info page
Stream: Internet Engineering Task Force (IETF)
RFC: 9110
STD: 97
Obsoletes: 2818, 7230, 7231, 7232, 7233,
Updates: 3864
Category: Standards Track
Published: June 2022
ISSN: 2070-1721
Authors: R. Fielding, Ed. Adobe
M. Notting Fastly
RFC 9110
HTTP Semantics
Abstract
The Hypertext Transfer Protocol (HTTP) is a stateless application-level protocol for distributed, collaborative, hypertext information systems. This document describes the overall architecture of HTTP, establishes common terminology, and defines aspects of the protocol that are shared by all versions. In this definition are core protocol elements, extensibility mechanisms, and the "http" and "https" Uniform Resource Identifier (URI) schemes.
This document updates RFC 3864 and obsoletes RFCs 2818, 7231, 7232, 7233, 7235, 7538, 7615, 7694, and portions of 7230.
Shared Understanding of HTTP Enables Communication
Memory
Application
20
8
Blocked requests
Size
16.8 kB
http/1.1 stylesheet asiaccs2024.sutd.ed (memory... 0..
http/1.1 stylesheet asiaccs2024.sutd.ed (memory... 0..
http/1.1 stylesheet asiaccs2024.sutd.ed (memory... 0..
http/1.1 stylesheet asiaccs2024.sutd.ed (memory... 0..
http/1.1 script asiaccs2024.sutd.ed (memory... 0..
http/1.1 script asiaccs2024.sutd.ed (memory... 0..
http/1.1 script asiaccs2024.sutd.ed (memory... 0..
h2 stylesheet asiaccs2024.sutd.ed (disk cac... 1...
h2 stylesheet asiaccs2024.sutd.ed (disk cac... 1...
62 requests
19.2 kB transferred
3.3 MB resources
Finish: 1.72 s
DOMContentLoaded: 1.08 s
Load: 1.11 s

Following the Specifications is Difficult

Following the Specifications is Difficult

Host of Troubles: Multiple Host Ambiguities in HTTP Implementations

Jianjun Chen[*†] chenjj13@mails.tsinghua.edu.cn

Jian Jiang[‡] jiangjian@berkeley.edu

Haixin Duan[*†] duanhx@tsinghua.edu.cn

Nicholas Weaver[‡§] nweaver@icsi.berkeley.edu

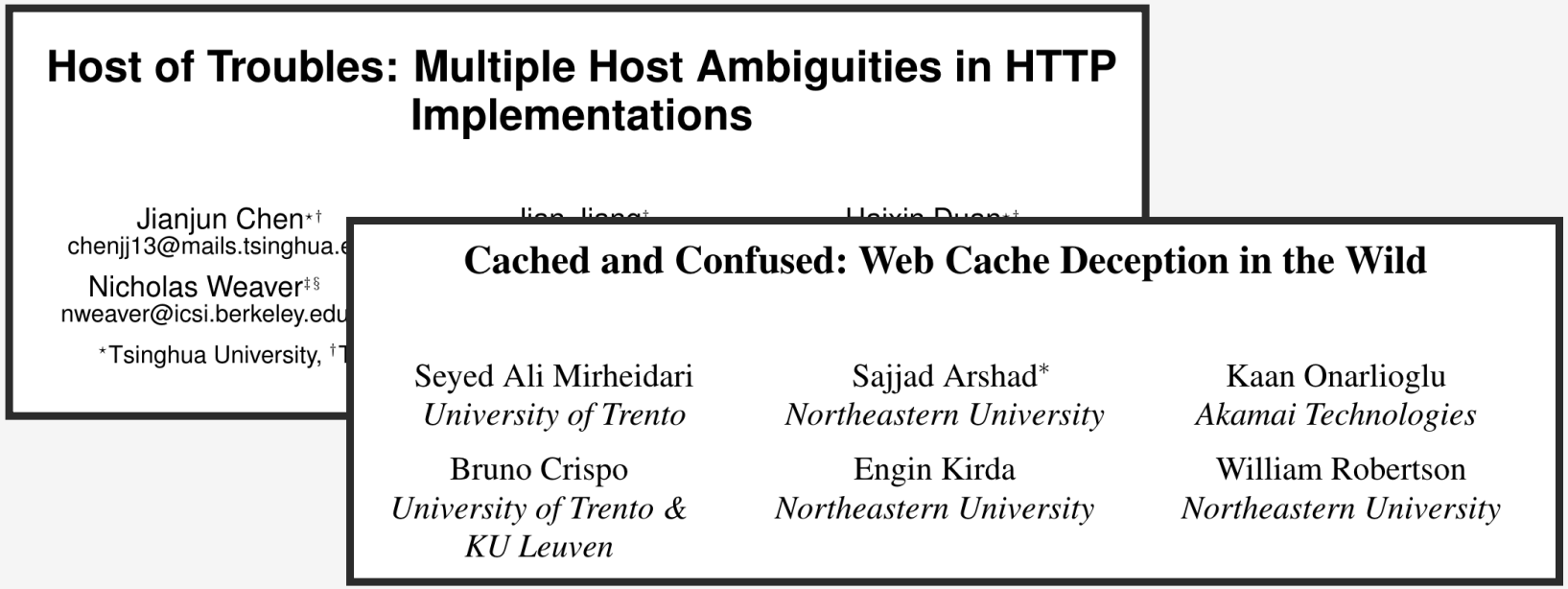
Tao Wan[¶] tao.wan@huawei.com

Vern Paxson[‡§] vern@berkeley.edu

[*]Tsinghua University, [†]Tsingua National Laboratory for Information Science and Technology
[‡]UC Berkeley, [§]ICSI, [¶]Huawei Canada

J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, and V. Paxson, “Host of Troubles: Multiple Host Ambiguities in HTTP Implementations,” CCS 2016

Following the Specifications is Difficult



Host of Troubles: Multiple Host Ambiguities in HTTP Implementations

Jianjun Chen*†
chenjj13@mails.tsinghua.e
Nicholas Weaver‡§
nweaver@icsi.berkeley.edu
*Tsinghua University, †T

Cached and Confused: Web Cache Deception in the Wild

Seyed Ali Mirheidari
University of Trento

Sajjad Arshad*
Northeastern University

Kaan Onarlioglu
Akamai Technologies

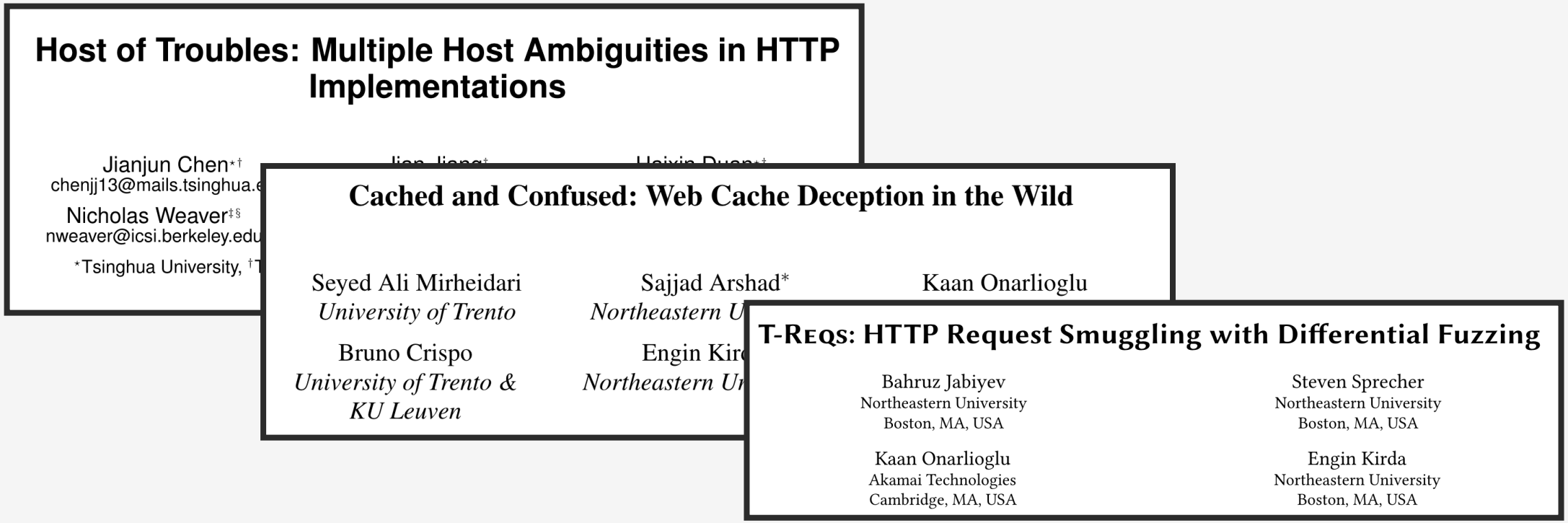
Bruno Crispo
University of Trento & KU Leuven

Engin Kirda
Northeastern University

William Robertson
Northeastern University

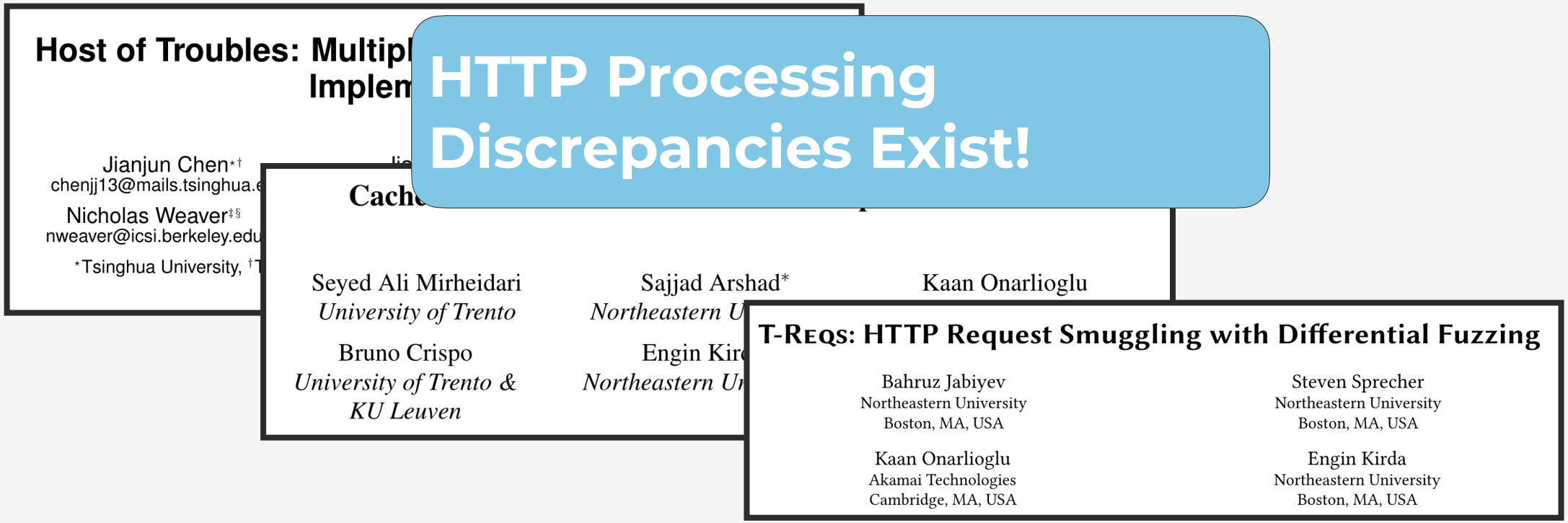
J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, and V. Paxson, "Host of Troubles: Multiple Host Ambiguities in HTTP Implementations," CCS 2016
S. A. Mirheidari, S. Arshad, K. Onarlioglu, B. Crispo, E. Kirda, and W. Robertson, "Cached and Confused: Web Cache Deception in the Wild," USENIX 2020

Following the Specifications is Difficult



J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, and V. Paxson, "Host of Troubles: Multiple Host Ambiguities in HTTP Implementations," CCS 2016
S. A. Mirheidari, S. Arshad, K. Onarlioglu, B. Crispo, E. Kirda, and W. Robertson, "Cached and Confused: Web Cache Deception in the Wild," USENIX 2020
B. Jabiyev, S. Sprecher, K. Onarlioglu, and E. Kirda, "T-Reqs: HTTP Request Smuggling with Differential Fuzzing," CCS 2021

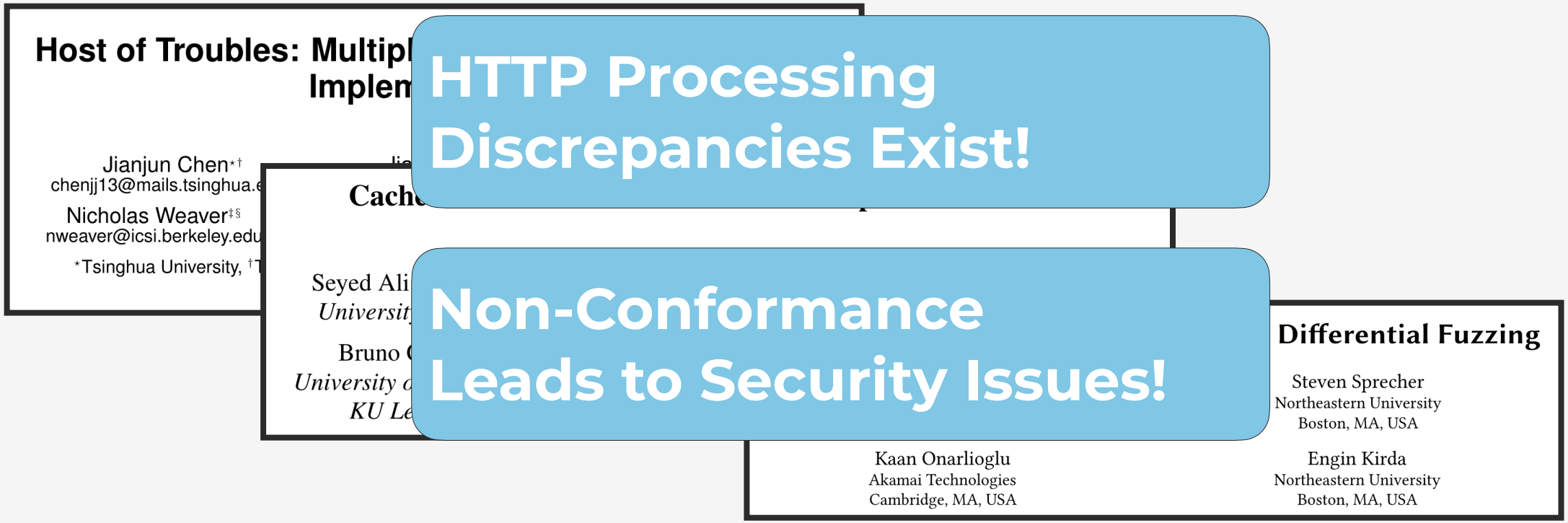
Following the Specifications is Difficult



HTTP Processing Discrepancies Exist!

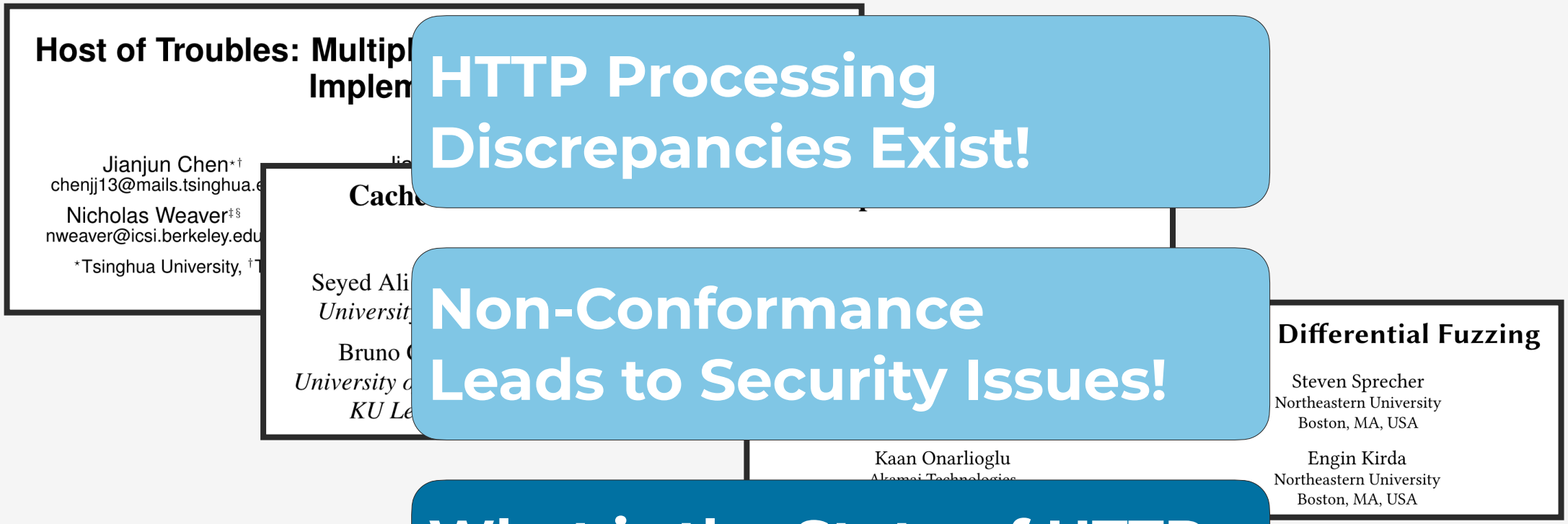
J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, and V. Paxson, "Host of Troubles: Multiple Host Ambiguities in HTTP Implementations," CCS 2016
S. A. Mirheidari, S. Arshad, K. Onarlioglu, B. Crispo, E. Kirda, and W. Robertson, "Cached and Confused: Web Cache Deception in the Wild," USENIX 2020
B. Jabiyev, S. Sprecher, K. Onarlioglu, and E. Kirda, "T-Reqs: HTTP Request Smuggling with Differential Fuzzing," CCS 2021

Following the Specifications is Difficult



J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, and V. Paxson, "Host of Troubles: Multiple Host Ambiguities in HTTP Implementations," CCS 2016
S. A. Mirheidari, S. Arshad, K. Onarlioglu, B. Crispo, E. Kirda, and W. Robertson, "Cached and Confused: Web Cache Deception in the Wild," USENIX 2020
B. Jabiyev, S. Sprecher, K. Onarlioglu, and E. Kirda, "T-Reqs: HTTP Request Smuggling with Differential Fuzzing," CCS 2021

Following the Specifications is Difficult



HTTP Processing Discrepancies Exist!

Non-Conformance Leads to Security Issues!

What is the State of HTTP Conformance in the Wild?

J. Chen, J. Jiang, H. Duan, N. Weaver, T. Wan, a
Multiple Host Ambiguities in HTTP Implementa
S. A. Mirheidari, S. Arshad, K. Onarlioglu, B. C
"Cached and Confused: Web Cache Deception in
B. Jabiyev, S. Sprecher, K. Onarlioglu, and E. Kirda, "T-Reqs: HTTP Request Smuggling with Differential Fuzzing," CCS 2021

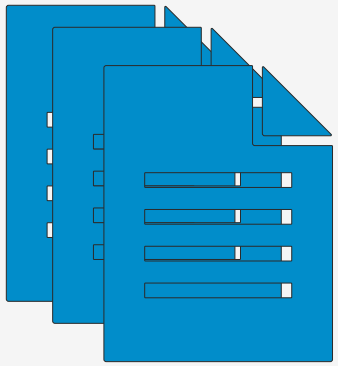
Measuring HTTP Conformance in the Wild

Measuring HTTP Conformance in the Wild

Rule Mining

Measuring HTTP Conformance in the Wild

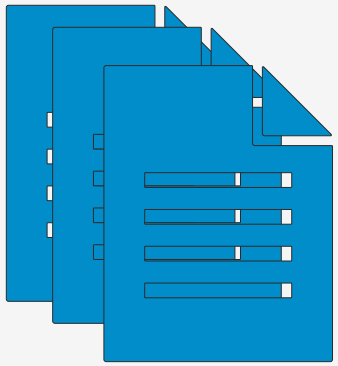
Rule Mining



HTTP Spec

Measuring HTTP Conformance in the Wild

Rule Mining

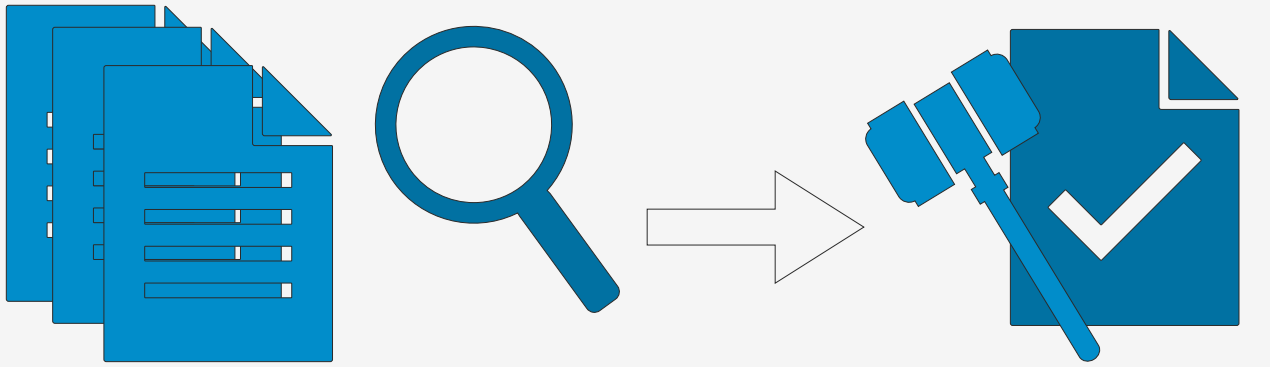


HTTP Spec

Rule Extraction

Measuring HTTP Conformance in the Wild

Rule Mining



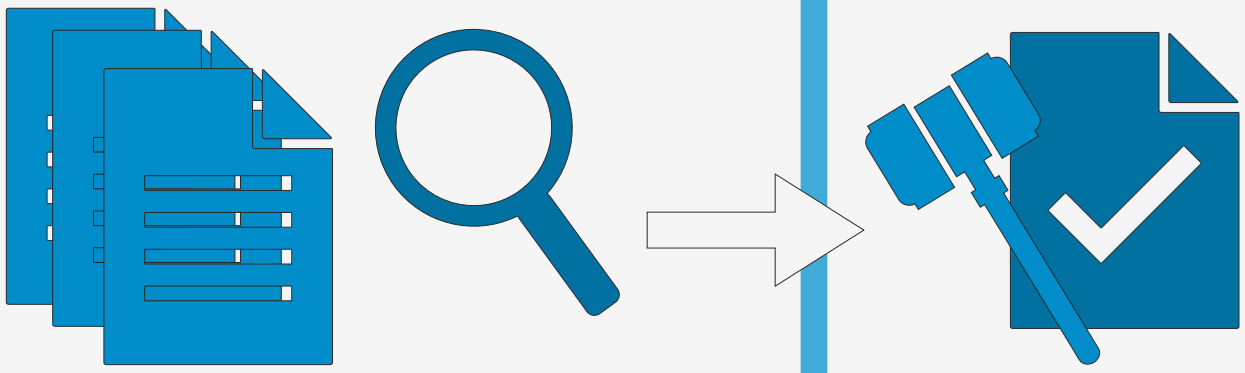
HTTP Spec

Rule Extraction

List of Rules

Measuring HTTP Conformance in the Wild

Rule Mining



HTTP Spec

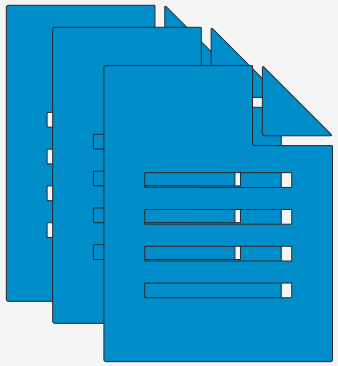
Rule Extraction

List of Rules

Conformance Testing

Measuring HTTP Conformance in the Wild

Rule Mining

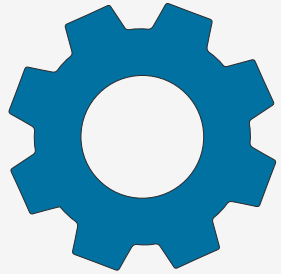


HTTP Spec

Rule Extraction

List of Rules

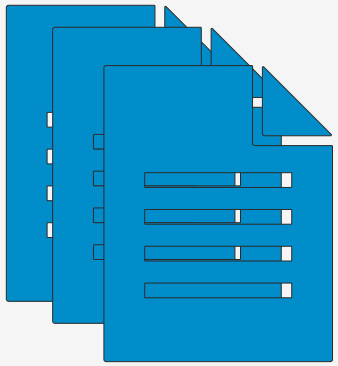
Conformance Testing



RuleBreaker

Measuring HTTP Conformance in the Wild

Rule Mining

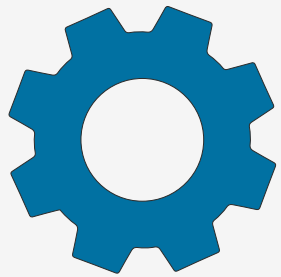


HTTP Spec

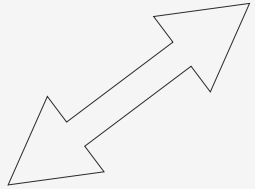
Rule Extraction

List of Rules

Conformance Testing



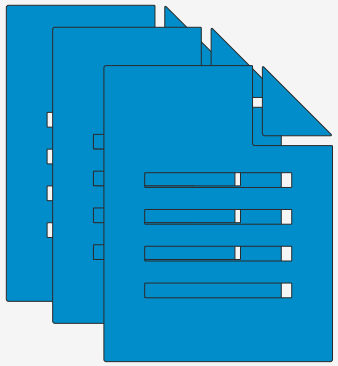
RuleBreaker



Web Server

Measuring HTTP Conformance in the Wild

Rule Mining

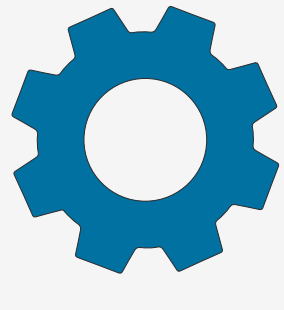


HTTP Spec

Rule Extraction

List of Rules

Conformance Testing



RuleBreaker

Web Server

Conformance Results

Extracting Rules

Extracting Rules

8.6. Content-Length

The "Content-Length" header field indicates the associated representation's data length as a decimal non-negative integer number of octets. When transferring a representation as content, Content-Length refers specifically to the amount of data enclosed so that it can be used to delimit framing (e.g., Section 6.2 of [HTTP/1.1]). In other cases, Content-Length indicates the selected representation's current length, which can be used by recipients to estimate transfer time or to compare with previously stored representations.

```
Content-Length = 1*DIGIT
```

An example is

```
Content-Length: 3495
```

A user agent **SHOULD** send Content-Length in a request when the method defines a meaning for enclosed content and it is not sending Transfer-Encoding. For example, a user agent normally sends Content-Length in a POST request even when the value is 0 (indicating empty content). A user agent **SHOULD NOT** send a Content-Length header field when the request message does not contain content and the method semantics do not anticipate such data.

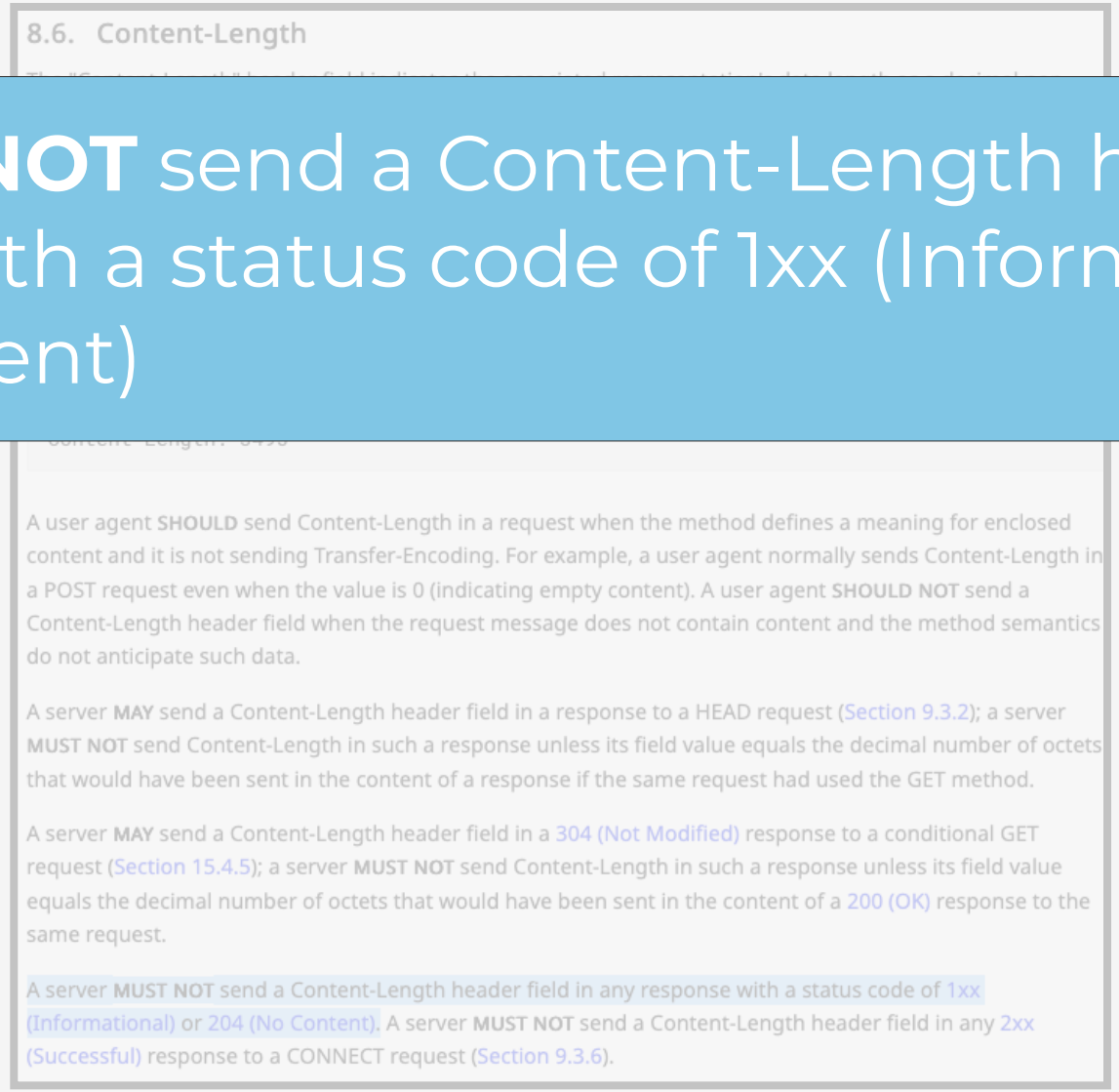
A server **MAY** send a Content-Length header field in a response to a HEAD request (Section 9.3.2); a server **MUST NOT** send Content-Length in such a response unless its field value equals the decimal number of octets that would have been sent in the content of a response if the same request had used the GET method.

A server **MAY** send a Content-Length header field in a 304 (Not Modified) response to a conditional GET request (Section 15.4.5); a server **MUST NOT** send Content-Length in such a response unless its field value equals the decimal number of octets that would have been sent in the content of a 200 (OK) response to the same request.

A server **MUST NOT** send a Content-Length header field in any response with a status code of 1xx (Informational) or 204 (No Content). A server **MUST NOT** send a Content-Length header field in any 2xx (Successful) response to a CONNECT request (Section 9.3.6).

Extracting Rules

A server **MUST NOT** send a Content-Length header field in any response with a status code of 1xx (Informational) or 204 (No Content)



8.6. Content-Length

A user agent **SHOULD** send Content-Length in a request when the method defines a meaning for enclosed content and it is not sending Transfer-Encoding. For example, a user agent normally sends Content-Length in a POST request even when the value is 0 (indicating empty content). A user agent **SHOULD NOT** send a Content-Length header field when the request message does not contain content and the method semantics do not anticipate such data.

A server **MAY** send a Content-Length header field in a response to a HEAD request (Section 9.3.2); a server **MUST NOT** send Content-Length in such a response unless its field value equals the decimal number of octets that would have been sent in the content of a response if the same request had used the GET method.

A server **MAY** send a Content-Length header field in a 304 (Not Modified) response to a conditional GET request (Section 15.4.5); a server **MUST NOT** send Content-Length in such a response unless its field value equals the decimal number of octets that would have been sent in the content of a 200 (OK) response to the same request.

A server **MUST NOT** send a Content-Length header field in any response with a status code of 1xx (Informational) or 204 (No Content). A server **MUST NOT** send a Content-Length header field in any 2xx (Successful) response to a CONNECT request (Section 9.3.6).

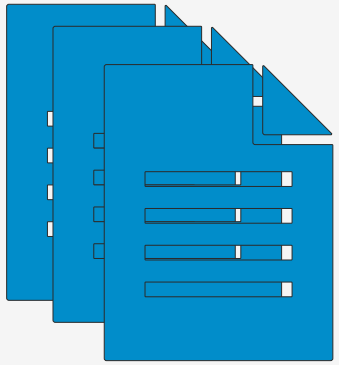
Extracting Rules

A server **MUST NOT** send a Content-Length header field in any response with a status code of 1xx (Informational) or 204 (No Content)

Rule: CL 1XX-204
Request: Any
Response: If (Code == 1XX or 204) and (CL exist) ⇨ Broken!

Test Environment

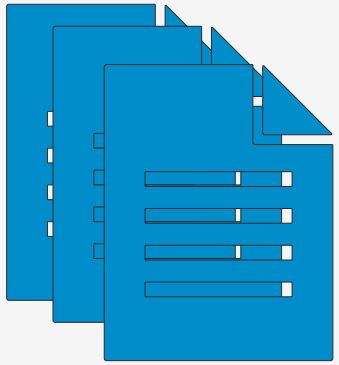
Test Environment



HTTP Specs

- 7 RFCs (IETF)
- 2 Living Standards (WHATWG)
- 3 Technical Reports (W3C)

Test Environment



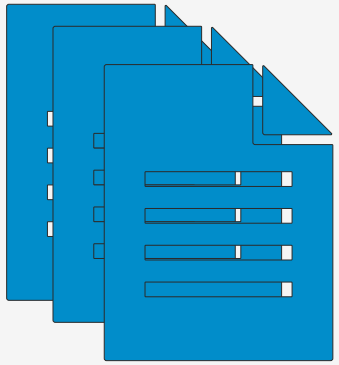
HTTP Specs

- 7 RFCs (IETF)
- 2 Living Standards (WHATWG)
- 3 Technical Reports (W3C)

List of Rules

- 106 extracted rules
- 55 security-relevant

Test Environment



HTTP Specs

- 7 RFCs (IETF)
- 2 Living Standards (WHATWG)
- 3 Technical Reports (W3C)

List of Rules

- 106 extracted rules
- 55 security-relevant

Web Servers

- 9 local servers (Apache, NGINX, ...)
- Top5K CrUX
- Random 5K Long-Tail

Results Overview

Results Overview

Local Servers

- 16 broken rules
- 6-10/Server

Results Overview

Local Servers

- 16 broken rules
- 6-10/Server

10K Web Sample

- 68 broken rules
- Average: 4.69/Host
- Max: 16
- Popular ≈ Long-Tail

Example Security-Relevant Violations

Example Security-Relevant Violations

Header Misconfigurations

- Duplicate Headers (465x)
- Syntax Violations (>250x)

Example Security-Relevant Violations

Header Misconfigurations

- Duplicate Headers (465x)
- Syntax Violations (>250x)

Common Reasons

- Header Confusion
- Encoding & Templating

Example Security-Relevant Violations

Header Misconfigurations

- Duplicate Headers (465x)
- Syntax Violations (>250x)

Dangerous Primitives

Common Reasons

- Header Confusion
- Encoding & Templating

Example Security-Relevant Violations

Header Misconfigurations

- Duplicate Headers (465x)
- Syntax Violations (>250x)

Dangerous Primitives

Common Reasons

- Header Confusion
- Encoding & Templating

10K Web Sample

- Content-Length 1XX-204 (55x)
- Content for 304 (46x)

Example Security-Relevant Violations

Header Misconfigurations

- Duplicate Headers (465x)
- Syntax Violations (>250x)

Dangerous Primitives

Common Reasons

- Header Confusion
- Encoding & Templating

10K Web Sample

- Content-Length 1XX-204 (55x)
- Content for 304 (46x)

Local Servers

- Host of Troubles (9/9)
- Illegal Chars (7/9)

Towards Improving HTTP Conformance

Towards Improving HTTP Conformance

- Conformance is bad
  - 75 rules broken
  - Average: 4.69/Host

Towards Improving HTTP Conformance

- Conformance is bad
  - 75 rules broken
  - Average: 4.69/Host
- Towards better conformance
  - No more: “Be conservative in what you send, be liberal in what you accept”
  - Shared test suite

Towards Improving HTTP Conformance

- Conformance is bad
  - 75 rules broken
  - Average: 4.69/Host
- Towards better conformance
  - No more: “Be conservative in what you send, be liberal in what you accept”
  - Shared test suite

Thanks for your attention!

https://github.com/cispa/http-conformance

jannis.rautenstrauch@cispa.de