

Pivot of the Week: Host Pairs

The pivot of the Week: Host Pairs

For more than ten years, RiskIQ has been crawling and absorbing the internet to define the web's identity and composition by fingerprinting each web component, connection, service, IP-connected device, and infrastructure to show how organizations and the attackers targeting them fit within it.

RisklQ's global collection network assembles, sorts, classifies, and monitors this internet data so analysts can leverage the relationships between these highly connected data sets to build out a thorough investigation. This process is the core of threat hunting and allows organizations to have a complete picture of attackers and their evolving techniques.

Host pairs is a data set derived from this process, unique to RisklQ PassiveTotal. Host pairs are two domains (a parent and a child) that shared a connection observed from a RisklQ web crawl, and its connections can range from a top-level redirect (HTTP 302) to something more complex, like an iframe or script source reference.

Host Pairs relies on knowing web site content, so it's likely to surface different values that other sources like passive DNS and SSL certificates do not. What makes this data set powerful is it gives researchers the ability to understand relationships between hosts based on details from visiting the actual page. Each connection has an observed first time and last time point that helps establish a time period for the pair of web properties and preserves a given relationship between two web properties that may no longer exist or remain valid.

Questions Host Pairs Can Answer

- Are there any additional suspicious hosts redirecting or being redirected from the indicator being investigated?
- What type of redirection is taking place?
- Where are users being redirected from/to?
- Are there any patterns with the redirections taking place?
- Does the number of pairs for a given property help in dictating if it's malicious or non-malicious?
- Have any of the connected artifacts been blacklisted?
- Have any of the connected artifacts been tagged as malicious?
- What is the source of the malware in this domain?
- Is this domain redirecting users to malicious content?
- Are resources pulling in CSS or images to set up infringement attacks?
- Are resources pulling in a Script to setup a magecart or skimming attack?

Host Pairs Causes in Action

RiskIQ tracks over 16,000 Host Pair causes. Many of them are HTML tags and attributes.

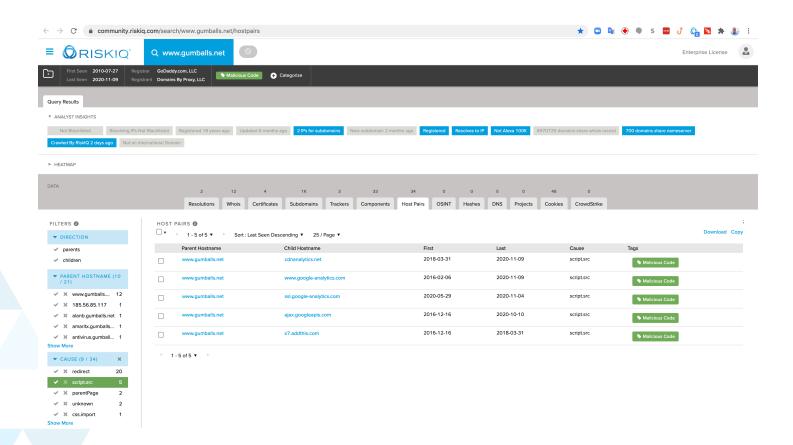
Img.src

What is an img.src cause?

The website is linking an external image that is hosted on another website.

https://community.riskiq.com/search/assets.nflxext.com/hostpairs

On the right side under **CAUSE** click the check icon next to img.src. This will filter the results to show just host pairs that were caused by img.src



script.src

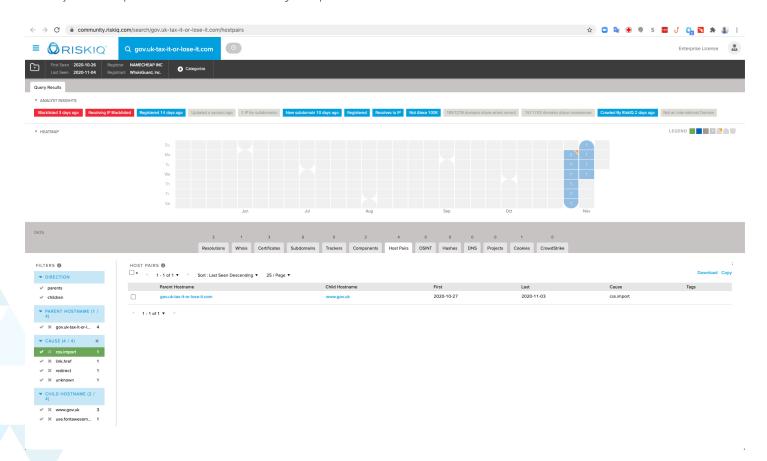
What does a Script cause mean?:

The src attribute specifies the domain where an external script file is pulled from to run.

Example: If you want to run the same JavaScript on several pages in a web site, you should create an external JavaScript file instead of writing the same script over and over again. But, malicious skimmers like Magecart will his is how threat actors run

https://community.riskig.com/search/www.gumballs.net/hostpairs

On the right side under **CAUSE** click the check icon next to script.src. This step will filter the results to show just host pairs that were caused by script.src



CSS import

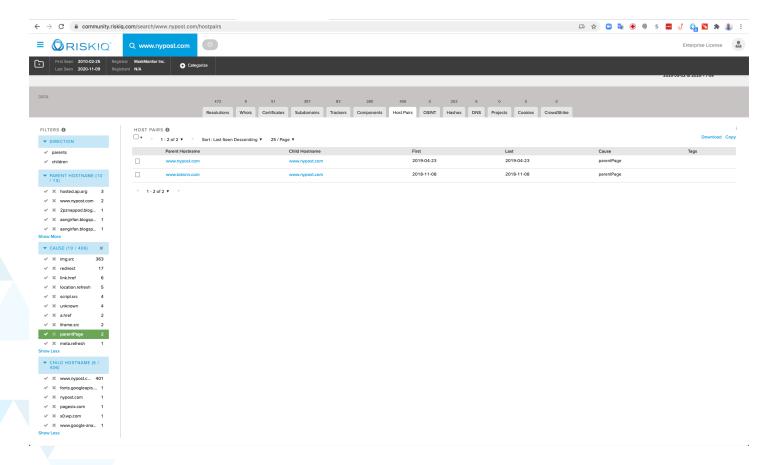
What does CSS import mean as a cause?

Cascading Style Sheets (CSS) is a stylesheet language used to describe the presentation of a document written in HTML or XML (including XML dialects such as SVG, MathML or XHTML). CSS describes how elements should be rendered on screen, on paper, in speech, or on other media.

Threat actors targeting phishing attacks will import the CSS from the domain they are targeting to duplicate the domain.

https://community.riskig.com/search/gov.uk-tax-it-or-lose-it.com/hostpairs

On the right side under **CAUSE** click the check icon next to CSS.import. This step will filter the results to show just host pairs that were caused by css.import.



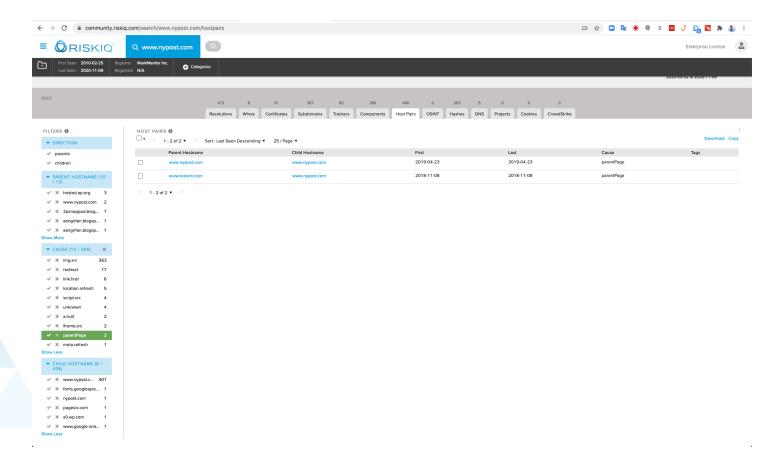
parentPage

What is a parent page cause?

A parent page is a top-level page of a website that may have child pages nested under it.

https://community.riskig.com/search/www.nypost.com/hostpairs

On the right side under **CAUSE** click the check icon next to 'parentpage.' This step will filter the results to show just host pairs that were caused by a parent page.



Redirect

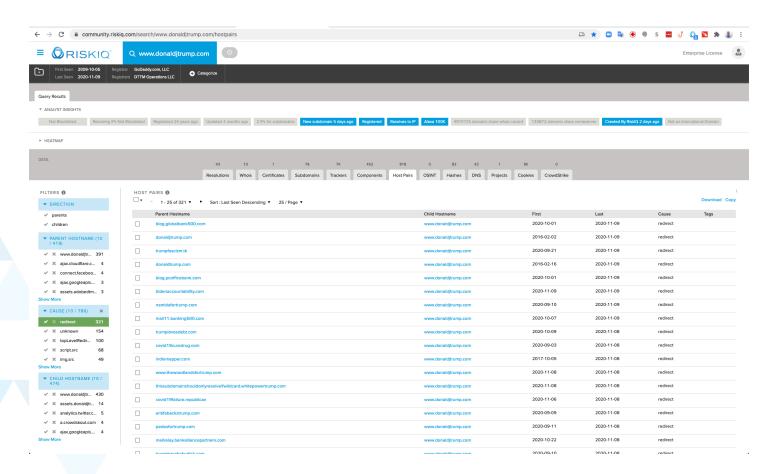
What does redirect cause mean?

Within HTTP response codes, the 3xx (Redirection) class of status code indicates that further action needs to be taken by the user agent to fulfill the request. If a Location header field is provided, the user agent MAY automatically redirect its request to the URI referenced by the Location field value, even if the specific status code is not understood.

During the Presidential election of 2020, we saw numerous redirects to both Presidential candidates' official domains. These redirects were to help or harm the candidates' reputation with voters.

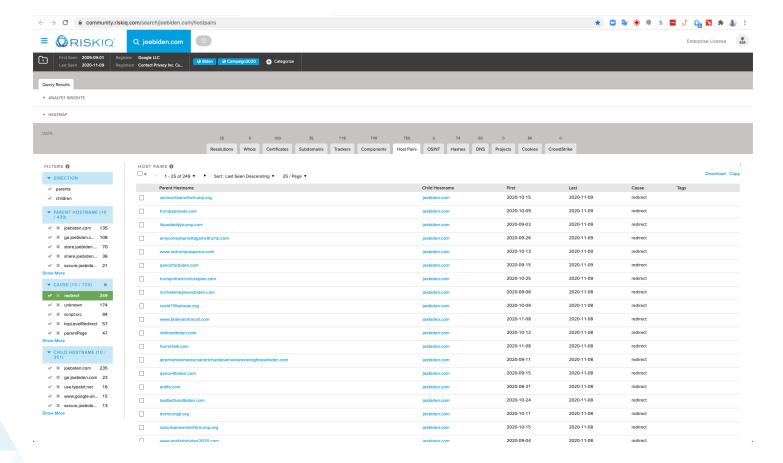
https://community.riskig.com/search/www.donalditrump.com/hostpairs

On the right side under **CAUSE** click the check icon next to redirect. This will filter the results to just show host pairs that were caused by a redirect.



https://community.riskiq.com/search/joebiden.com/hostpairs

On the right side under **CAUSE** click the check icon next to redirect. This will filter the results to show just host pairs that were caused by a redirect.



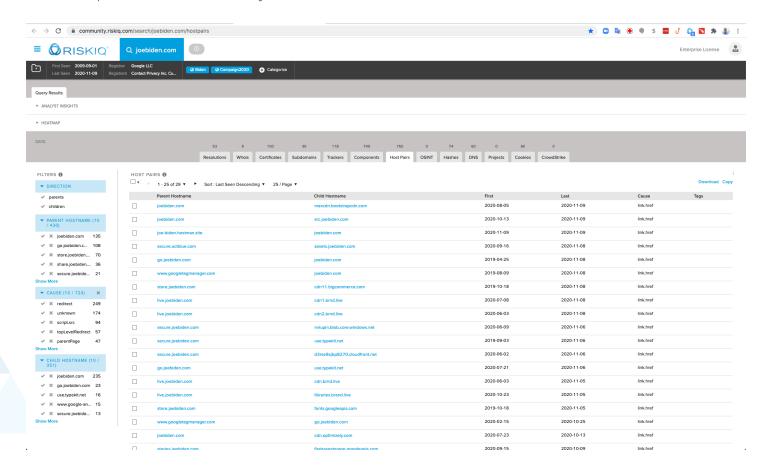
Link.href

What does a link.href mean as a cause?

The href attribute specifies the URL of the page the link goes to.

https://community.riskiq.com/search/joebiden.com/hostpairs

On the right side under **CAUSE** click the check icon next to link.href. This will filter the results to just show host pairs that were caused by a link.href.



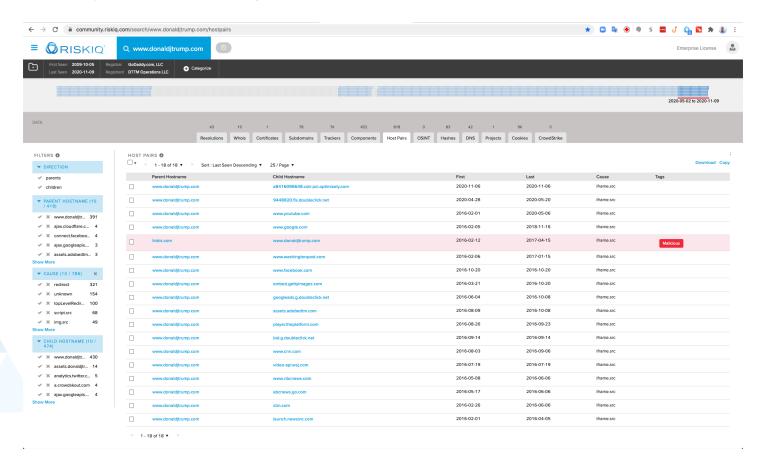
Iframe.src

What is an iframe.src?

The src attribute specifies the address of the document to embed in the iframe. Within PassiveTotal the domain and relationship is seen but not the full path.

https://community.riskiq.com/search/www.donaldjtrump.com/hostpairs

On the right side under **CAUSE** click the check icon next to iframe.src. This will filter the results to just show host pairs that were caused by iframe.src.



Xmlhttprequest

What is an xmlhttprequest cause?

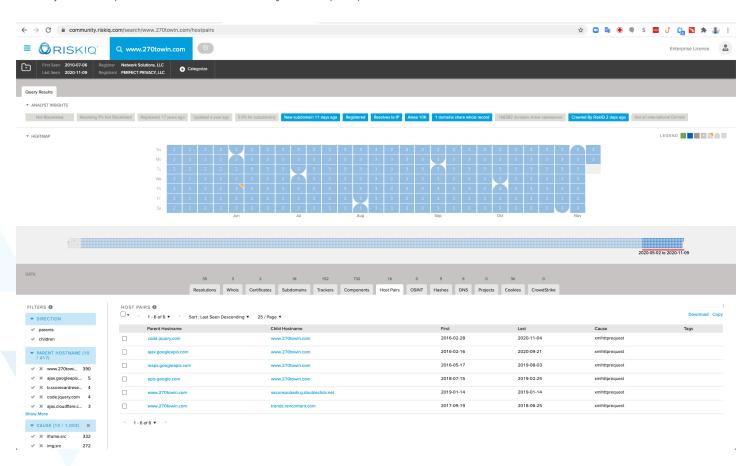
The XMLHttpRequest object can be used to request data from a web server.

The XMLHttpRequest object can be used to:

- Update a web page without reloading the page
- Request data from a server after the page has loaded
- Receive data from a server after the page has loaded
- Send data to a server in the background

https://community.riskiq.com/search/www.270towin.com/hostpairs

On the right side under **CAUSE** click the check icon next to xmlhttprequest. This will filter the results to show just host pairs that were caused by xmlhttprequest.



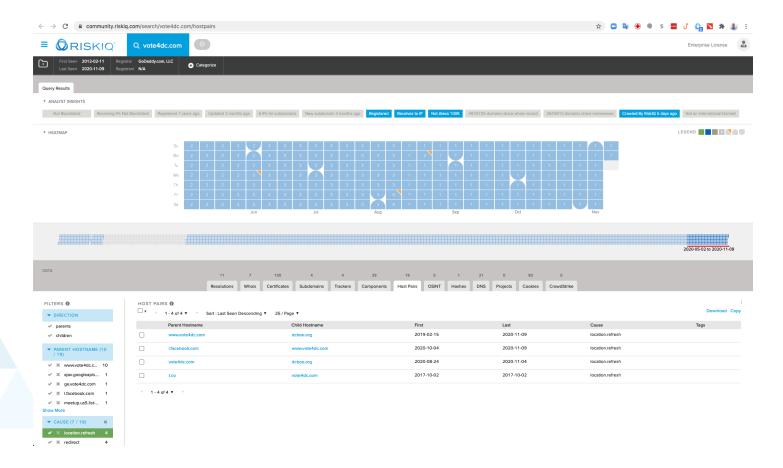
Location.refresh

What does a location.refresh cause mean?

The Location.reload() method reloads the current URL, like the Refresh button in a web browser.

https://community.riskig.com/search/vote4dc.com/hostpairs

On the right side under **CAUSE** click the check icon next to location.refresh. This will filter the results to show just host pairs that were caused by location.refresh.



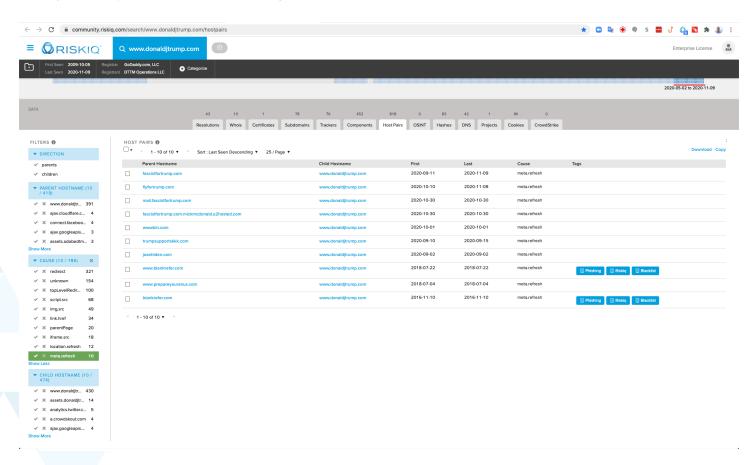
Meta.refresh

What is a meta.refresh cause?

Meta refresh is a method of instructing a web browser to automatically refresh the current web page or frame after a given time interval, using an HTML meta element with the http-equiv parameter set to "refresh" and a content parameter giving the time interval in seconds. It is also possible to instruct the browser to fetch a different URL when the page is refreshed by including the alternative URL in the content parameter. By setting the refresh time interval to zero (or a very low value), a meta refresh can be used as a URL redirection method.

https://community.riskig.com/search/www.donalditrump.com/hostpairs

On the right side under **CAUSE** click the check icon next to meta.refresh. This will filter the results to show just host pairs that were caused by meta.refresh.

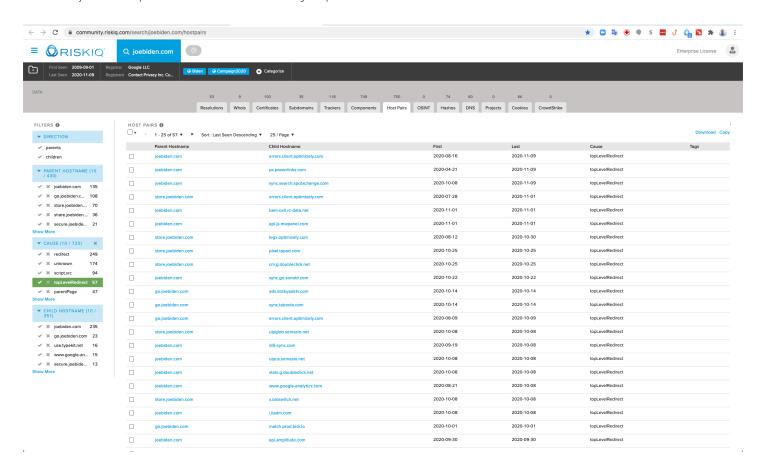


topLevelRedirect

This is a redirection to the top-level domain. For example, if you owned two domains example.com and also exmaple.com (common misspelling), you could do a top-level redirect of exmaple.com to go to example.com. This would help users that typed the domain incorrectly to get to the appropriate domain.

https://community.riskiq.com/search/joebiden.com/hostpairs

On the right side under **CAUSE** click the check icon next to topLevelRedirect. This will filter the results to show just host pairs that were caused by topLevelRedirect.





RiskIQ, Inc.

22 Battery Street, 10th Floor San Francisco, CA. 94111



L 1 888.415.4447

Learn more at riskiq.com

Copyright © 2020 RisklQ, Inc. RisklQ, the RisklQ logo and RisklQ family of marks are registered trademarks or trademarks of RisklQ, Inc. in the United States and other countries. Other trademarks mentioned herein may be trademarks of RisklQ or other companies. 08_20