# TARSAN API Protocol

Version 1.5 – August 11th, 2015

by: Gunarto Sindoro Njoo (532711)

## General URL Format

http://<Server name>:<Server Port>/<Project Name>/<Entry Point>/<Path>?service=<Service name>&jsonPara=<JSON parameter>

* Server name: Host of the API server (e.g., 140.96.29.153)
* Server port: Port of the API server (through Tomcat) (e.g., 8080)
* Project name: The name of the project (i.e., TARSAN)
* Entry point: Entry point name, defined in web.xml (src/main/webapp/WEB-INF/web.xml)
* Path: name of the entry point path, defined in Tarsan.java (@path)
* Service name: name of the service that is used (e.g., login, createUser, etc.)
* JSON parameter: parameter for the invoked service

## JSON Status codes

Update July 21, 2015 – 01.17 PM

|  |  |  |
| --- | --- | --- |
| Status code | Status name | Remarks |
| 000 | CODE\_EXCEPTION | When exceptions occurs |
| 200 | CODE\_OK | When operation success |
| 400 | CODE\_ERROR | Generally, when error happened |
| 401 | CODE\_UNAUTHORIZED | Ask for API operation without correct session |
| 402 | CODE\_NOT\_EXIST | Not found in database, or other sources |
| 403 | CODE\_FORBIDDEN | Couldn’t access the API service because operation or parameter is invalid |
| 405 | CODE\_NOT\_ALLOWED | Couldn’t access the API service because service hasn’t been implemented yet |
| 520 | CODE\_UNKNOWN\_ERR | Unknown error occurred |

## BoUser.java

### LOGIN

Update July 21, 2015 – 01.17 PM

Authentication process, query account name and password from database.  
Error will happen if incorrect password is given or account doesn’t exist in the database.

PS: Login status in Squid will be stored for 2 hours, so if we use Proxy mode, then for 2 hours, the session for the same user will be regarded as the same session.

**Input**

* account: account username
* password: account password

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 000 | Exception |  |
| 200 | Login success | [{"status":"200"},{"login":"1plkw46n43f6e8jlofvdq88n{message=User lalala successfully logged in, tmp\_key=aa811018e6f75b854be9d9c9e51efe66}] |
| 402 | Account not exist | [{"status":"402"},{"login":null},{"message":"Account does not exist"}] |
| 403 | Wrong password | [{"status":"403"},{"login":null},{"message":"Invalid password"}] |

status: Represent the result of the query.

login: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

tmp\_key: generated key for identifying user’s session.

message: Informative message for human.

**Test case:**

|  |
| --- |
| Success   * [http://140.96.29.153:8080/TARSAN/service/main?service=login&jsonPara=[("account":"lalala"),("password":"lalala")](http://140.96.29.153:8080/TARSAN/service/main?service=login&jsonPara=%5b(%22account%22:%22lalala%22),(%22password%22:%22lalala%22))]   Failed   * [http://140.96.29.153:8080/TARSAN/service/main?service=login&jsonPara=[("account":"lalala1"),("password":"lalala1")](http://localhost:8080/TARSAN/service/main?service=login&jsonPara=%5b(%22account%22:%22lalala1%22),(%22password%22:%22lalala1%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=login&jsonPara=[("account":"lalala"),("password":"lalala1")](http://localhost:8080/TARSAN/service/main?service=login&jsonPara=%5b(%22account%22:%22lalala%22),(%22password%22:%22lalala1%22))] |

### REGISTER

Update July 21, 2015 – 01.17 PM

Registration of new user by using two parameters: account and password.  
Error will happen if format is incorrect (e.g., length of password is too short or too long) or account id has already existed in the database.

**Input**

* account: account username
* password: account password

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 000 | Exception |  |
| 200 | Create new user successfully | [{"status":"200"},{"createUser":"1plkw46n43f6e8jlofvdq88n"},{"message":"User abc123 is successfully created"}] |
| 400 | Unable to create the new user | [{"status":"400"},{"createUser":null},{"message":"Account with name [abc123] is already registered"}] |
| 403 | Input format is incorrect (password length is more than 128 chars or less than 3 chars) | [{"status":"403"},{"createUser":null},{"message":"Input format or length incorrect"}] |

status: Represent the result of the query.

createUser: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

**Test case:**

|  |
| --- |
| [http://140.96.29.153:8080/TARSAN/service/main?service=createUser&jsonPara=[("account":"lalala"),("password":"lalala")](http://localhost:8080/TARSAN/service/main?service=createUser&jsonPara=%5b(%22account%22:%22lalala%22),(%22password%22:%22lalala%22))] |

### LOG USER EVENT

Update September 03, 2015 – 11.18 PM

Logging user events from his/her browsing history. Two modes are offered: Normal and ICAP.

**Input**

* url
* referrer: referrer website
* userId or username (userId for browser extension and username for ICAP/Proxy)
* patternId: pattern Id that is match
* ip: ip address
* productSearchFlag: indicate if the log reflect a “product search mode”
* parent Id: referrer id
* root Id: starter of the session’s id

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 000 | Exception |  |
| 200 | Create new user log successfully | [{"status":"200"},{"logUserEvent":"6fde82ad29a1d9396094729e1bbc3790"},{"message":"user event has been logged","url":"http://buy.mi.com/tw/search\_mi note pro","domain":"buy.mi.com"}] |

status: Represent the result of the query.

logUserEvent: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

**Test case:**

|  |
| --- |
| Normal  [http://140.96.29.153:8080/TARSAN/service/main?service=logUserEvent&jsonPara=[("url":"http%3A%2F%2Fglobal%2Erakuten%2Ecom%2Fen%2Fsearch%2F%3Ftl%3D%26k%3Dmi%2Bnote%2Bpro"),("referrer":"localhost"),("userId":1),("patternId":1),("ip":"testing\_ip"),("productSearchFlag":true)](http://140.96.29.153:8080/TARSAN/service/main?service=logUserEvent&jsonPara=%5b(%22url%22:%22http%3A%2F%2Fglobal%2Erakuten%2Ecom%2Fen%2Fsearch%2F%3Ftl%3D%26k%3Dmi%2Bnote%2Bpro%22),(%22referrer%22:%22localhost%22),(%22userId%22:1),(%22patternId%22:1),(%22ip%22:%22testing_ip%22),(%22productSearchFlag%22:true)),("parentId ":0),("rootId":0)]  ICAP  [http://140.96.29.153:8080/TARSAN/service/main?service=logUserEventICAP&jsonPara=[(%22url%22:%22http%3A%2F%2Fbuy%2Emi%2Ecom%2Ftw%2Fsearch\_mi%2520note%2520pro%22),(%22referrer%22:%22localhost%22),(%22username%22:%22lalala%22),(%22patternId%22:1),(%22ip%22:%22testing\_ip%22),(%22productSearchFlag%22:true)](http://140.96.29.153:8080/TARSAN/service/main?service=logUserEventICAP&jsonPara=%5b(%22url%22:%22http%3A%2F%2Fbuy%2Emi%2Ecom%2Ftw%2Fsearch_mi%2520note%2520pro%22),(%22referrer%22:%22localhost%22),(%22username%22:%22lalala%22),(%22patternId%22:1),(%22ip%22:%22testing_ip%22),(%22productSearchFlag%22:true)),("parentId":0),("rootId":0)] |

## BoPattern.java

### RETRIEVE ADS URL

Update September 03, 2015 – 11.19 PM

|  |
| --- |
| Normal  [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206"),(“referrer”:”localhost”),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206%22),(),("parentId":0),("rootId":0)]  ICAP  [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLinkICAP&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206"),(“referrer”:”localhost”),("username":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLinkICAP&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206%22),(),("parentId ":0),("rootId":0)] |

Retrieve ads URL by giving a URL and domain name to the API.

**Input**

* url: Full URL of the website (e.g., <http://www.apple.com/search/?section=global&geo=us&q=iphone%206>)
* referrer: Previous website that is accessed before this url.
* temp\_key: user’s temp key which is obtained when user logged in. [For ICAP, it uses username instead of temp\_key, and it is converted in PatternFacade.java]
* parent Id: referrer id
* root Id: starter of the session’s id

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 000 | Exception |  |
| 200 | Ads delivery success | [{"status":"200"},{"getAdsLink":"1nxek9kaiqzh31sum41xvn2kux"},{"message":"ads url found","ads\_url":"http://140.96.29.33/tarsan/test.php?kw=iphone 6&domain=www.apple.com"}] |
| 401 | Temp key is incorrect | [{"status":"401"},{"getAdsLink":""},{"message":"invalid key"}] |
| 402 | Query pattern is different | [{"status":"402"},{"getAdsLink":""},{"message":"keyword couldn't be extracted","type":1,"url":"http://www.apple.com/search/?section=global&geo=us&qq=iphone 6","domain":"www.apple.com","url\_path":"/search/","parameters":["section=global","geo=us","qq=iphone 6"]}] |
| 402 | Query pattern is not in the database | [{"status":"402"},{"getAdsLink":""},{"message":"search pattern is not exist on the database or is expired","url":"http://www.apple.com/macbook-air/specs.html","domain":"www.apple.com","url\_path":"/macbook-air/","parameters":[]}] |
| 402 | Query type is not supported yet | [{"status":"402"},{"getAdsLink":""},{"message":"pattern type is not supported","type":2,"url":"http://www.crutchfield.com/shopsearch/dock.html","domain":"www.crutchfield.com","url\_path":"/shopsearch/","parameters":[]}] |
| 520 | Error in indexing (for debugging purpose) |  |

status: Represent the result of the query.

getAdsLink: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

ads\_url: advertisement URL.

url: original URL from input.

domain: hostname.

type, url\_path and parameters: parameters for search pattern matching.

**Test case:**

|  |
| --- |
| Sucess   * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206"),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206%22),(%22tempKey%22:%22test_123%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.rakuten.com/sr/searchresults.aspx%3Fqu=nokia"),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.rakuten.com/sr/searchresults.aspx%3Fqu=nokia%22),(%22tempKey%22:%22test_123%22))]   Failed   * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/macbook-air/specs.html"),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/macbook-air/specs.html%22),(%22tempKey%22:%22test_123%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26qq=iphone%206"),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26qq=iphone%206%22),(%22tempKey%22:%22test_123%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.crutchfield.com/shopsearch/dock.html"),("tempKey":"test\_123")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.crutchfield.com/shopsearch/dock.html%22),(%22tempKey%22:%22test_123%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206"),("tempKey":"test\_1234")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206%22),(%22tempKey%22:%22test_1234%22))] * [http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=[("url":"http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206"),("tempKey":"lalala")](http://140.96.29.153:8080/TARSAN/service/main?service=getAdsLink&jsonPara=%5b(%22url%22:%22http://www.apple.com/search/%3Fsection=global%26geo=us%26q=iphone%206%22),(%22tempKey%22:%22lalala%22))] |

### CREATE PATTERN

Update August 18, 2015 – 03.02 PM

|  |
| --- |
| Type 1  [http://140.96.29.153:8080/TARSAN/service/main?service=createPattern&jsonPara=[("url\_host":"www.testing.com"),("url\_path":"/search/"),("type":1),("signature":"q="),("webType":99)](http://140.96.29.153:8080/TARSAN/service/main?service=createPattern&jsonPara=%5b(%22url_host%22:%22www.testing.com%22),(%22url_path%22:%22/search/%22),(%22type%22:1),(%22signature%22:%22q=%22),(%22webType%22:99))] |

Create new pattern by using a few parameters: hostname, path, type of search query, signature, and EC website flag. Error will happen if pattern has already existed in the database.

**Input**

* url\_host: website hostname (e.g., [www.google.com](http://www.google.com), [www.apple.com](http://www.apple.com), etc.)
* url\_path: Path of the search query
* type: Search query type
  + Type 1: Normal – using signature
  + Type 2: Path – using url path
  + Type 99: Unknown
* webType: Type of website
  + Type 1: E-Commerce
  + Type 2: Search Engine
  + Type 3: News
  + Type 99: Unknown
* signature: Parameter name which indicate the position of the keyword that wants to be extracted.
* EC\_flag: indicating if it is a whitelist or a blacklist.

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 200 | New search query pattern is created | [{"status":"200"},{"createSearchPattern":"1wptztf54h6td19ii75k6y5brd"},{"message":"new search pattern is successfully created","host":"www.rakuten.com","type":3,"signature":"qu=","is\_EC":true}] |
| 400 | Unable to create the new search query pattern | [{"status":"400"},{"createSearchPattern":"1wptztf54h6td19ii75k6y5brd"},{"message":"search pattern has already existed on the database","url\_host":"www.rakuten.com","url\_path":null,"type":3,"signature":"qu="}] |

status: Represent the result of the query.

createSearchPattern: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

url\_host: Hostname of the website.

url\_path: Path of the search query.

type: Search query type.

signature: Parameter name which indicate the position of the keyword that wants to be extracted.

### GENERATE PATTERN LIST

Update July 21, 2015 – 01.44 PM

|  |
| --- |
| <http://140.96.29.153:8080/TARSAN/service/main?service=generatePatternList&jsonPara>=[] |

Generate pattern list from the database, in JSON Format.  
Error will happen if format is incorrect (e.g., length of string is too short or too long) or account id has already existed in the database.

**Input:** None

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 000 | Exception |  |
| 200 | Get the list of patterns from database | [{"status":"200"},{"generatePatternList":"hkqa1nrv70c2azzdx3bs33cp"},{"message":"Pattern list generated","size":"3","pattern":"[{\"patternId\":1,\"urlHost\":\"www.apple.com\",\"urlPath\":\"/search/\",\"type\":1,\"signature\":\"q=\",\"createdDateTime\":1437380506165,\"lastAccess\":null,\"isExpired\":false,\"isDeleted\":false,\"isEcWeb\":true}, {\"patternId\":2,\"urlHost\":\"www.crutchfield.com\",\"urlPath\":\"/shopsearch/\",\"type\":2,\"signature\":null,\"createdDateTime\":1437380510292,\"lastAccess\":null,\"isExpired\":false,\"isDeleted\":false,\"isEcWeb\":true}, {\"patternId\":3,\"urlHost\":\"www.rakuten.com\",\"urlPath\":null,\"type\":3,\"signature\":\"qu=\",\"createdDateTime\":1437380514119,\"lastAccess\":null,\"isExpired\":false,\"isDeleted\":false,\"isEcWeb\":true}]"}] |

status: Represent the result of the query.

generatePatternList: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

size: number of element inside the pattern array.

pattern: an array of pattern in the database.

**Pattern structure:**

++ patternId (long)

++ urlHost (String)

++ urlPath (String)

++ type (int)

++ signature (String)

++ createdDateTime (UNIX timestamp - Long)

++ lastAccess (UNIX timestamp - Long)

++ isExpired (Boolean)

++ isDeleted (Boolean)

++ isEcWeb (Boolean)

### LOG ADS

This function is not a service.

Update August 20, 2015 – 10.05 PM

Log\_ads is a function that helps the process of retrieving ads URL. While ads url retrieval only retrieves the URL, it needs another function to log the ads that is retrieved.

**Input**

* userId
* patternId
* ads\_url: URL of the ads. In the testing phase, it is generated from JSP.
* keyword
* identifier: Identify where the request come from (IP / Hostname / other).

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 200 | New search query pattern is created | [{status=200}, {logAds=testlogAds123}, {message=ad has been logged, ad=2, ad2user=23, temp\_key=33cc1c428f3e973485dc8567c2649a81}] |
| 403 | Violates ads showing policy | [{status=403}, {logAds=null}, {message=Violates ad showing policy, adId=2, url=http://140.96.29.153:8080/TARSAN/index.jsp?kw=iphone%25206&domain=www.apple.com}] |

status: Represent the result of the query.

logAds: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

ad: Advertisement id.

ad2user: Corresponding ad to user id

temp\_key: user’s temporal key (session)

url: Ads url.

## BoAds.java

### CREATE ADS

Update August 6, 2015 – 02.54 PM

|  |
| --- |
| http://140.96.29.153:8080/TARSAN/service/main?service=createAds&jsonPara=[(%22keywords%22:%22iphone%22),(%22url%22:%22www.apple.12345.com%22),(%22img%22:null),(%22weight%22:1.0)] |

Create new ads using keywords, URL path, image path, and weight. Error will happen if ad has already existed in the database.

**Input**

* keywords: keyword used in this ads
* url: Full URL of the ads.
* img: URL of the image that corresponds to the ads.
* weight: Weight of corresponding ads.

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 200 | New ads is created | [{status=200}, {create=1testCreateAds}, {message=new ads is successfully created, ad\_id=1103}] |
| 400 | Unable to create the new ads | [{status=400}, {create=null}, {message=ads has already existed on the database}] |

status: Represent the result of the query.

create: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.

### TOUCH

Update August 27, 2015 – 10.44 AM

|  |
| --- |
| [http://140.96.29.153:8080/TARSAN/service/main?service=touch&jsonPara=[(%22adsId%22:99),(%22ad2userId%22:209),(%22tempkey%22:%22test\_123%22),(%22timestamp%22:123456789),(%22identifier%22:%22localhost%22)](http://140.96.29.153:8080/TARSAN/service/main?service=touch&jsonPara=%5b(%22adsId%22:99),(%22ad2userId%22:209),(%22tempkey%22:%22test_123%22),(%22timestamp%22:123456789),(%22identifier%22:%22localhost%22))] |

Handle the event of ads click. If corresponding ad is clicked, then it updates click status of ad2user and ads’ number of click is increased by one.

**Input**

* keywords: keyword used in this ads
* url: Full URL of the ads.
* img: URL of the image that corresponds to the ads.
* weight: Weight of corresponding ads.

**JSON Message**

|  |  |  |
| --- | --- | --- |
| Status | Details | Example of output |
| 200 | Ad click event is success | [{"status":"200"},{"touch":"1tiyoesyxp68i1ewgtagfus1k7"},{"message":"ads is successfully touched","ad\_id":1,"click":102,"identifier":"S4kSYhuGE6XbUrg2gNxDTuB8NLe0EhfLqYGoPWUSH851yv6x4CfGiGPSswVj9Ksi9taELs5brVh2qDc82VWZhlJAItnU29kgpdKm0557927125"}] |
| 402 | Ad is not exist | [{"status":"402"},{"touch":null},{"message":"ads doesn't exist on the database"}] |
| 402 | No ad2user correspond to this ads | [{status=402}, {touch=null}, {message=No ad2user correspond to this ads}] |
| 403 | Not allowed to add more click rate (to prevent click fraud) | [{"status":"403"},{"touch":null},{"message":"Not allowed to add more click rate"}] |

status: Represent the result of the query.

create: Success: Session ID (random generated number).

Failures: <NULL>.

Unhandled exception: 0.

message: Informative message for human.