



## **DIMES REST Web Services – Returned Data Formats**

### 1. General:

1. DIMES REST Web Services were built in order to allow downloading DIMES topologies directly from the DIMES Data Base.
2. The service is opened to every one with no need of identification or authentication.
3. The topologies are built in weekly resolution when every week starts in Monday 00:00:00. First week: 37 / 2004.

#### 4. URL Pattern:

**[http://www.netdimes.org/rest/topologies/{file-type}/{year}/{week}\[?offset=x&limit=y\]](http://www.netdimes.org/rest/topologies/{file-type}/{year}/{week}[?offset=x&limit=y])**

1. file-type = one of the available topology files: nodes, edges, asnodes, asedges, citynodes, cityedges or routers (see further explanations in next sections).
  2. year = year of the asked topology, from 2004 to current.
  3. week = week of the asked topology, between 1 to 52.
  4. Optional: adding limit and offset parameters to the URL in case you want to download only part of the files in one transaction. If you add limit with no offset the offset is zero (the beginning of the topology).
  5. With no limit & offset parameters the full file will be downloaded in one transaction – be aware it can take a lot of space (up to hundreds of MB – depends of the size of the topology).
  6. Due to size and memory limits edges or nodes topologies can be downloaded only in transactions of up to 500000 rows. Topologies which are smaller than that can be downloaded in one transaction. Topologies which are bigger than that need to be downloaded in chunks of up to 500000 rows in each transaction.
5. For checking number or rows in a topology you can call the specific topology URL with '/count' as a postfix to get the rows count of this topology.
- For example: calling URL <http://www.netdimes.org/rest/topologies/edges/2008/52/count> will return:
- ```
<rows-count table-name="EdgesTbl_2008_week_52" table-type = "EDGES" >  
<count>1601314</count>  
</rows-count>
```
6. All tags in all kinds of XML result files are written in small letters.
  7. For further questions please contact: [support@netdimes.org](mailto:support@netdimes.org)



## 2. AS Nodes Data:

1. Represents nodes in Autonomous Systems (Ases) level.
2. Each row (AS node) includes this data:
  1. as-number – The number of the AS which this node belongs to.
  2. agent-who-discovered – The inner index of the DIMES agent who discovered this node (can be used for further experiments).
  3. date-of-discovery – The first date in the week in which this node was seen. Pattern: “yyyy-mm-dd hh:mm:ss”.
  4. date-of-validation – The last date in the week in which this node was seen. Pattern: “yyyy-mm-dd hh:mm:ss”.

### 3. Example of data

(URL: <http://www.netdimes.org/rest/topologies/asnodes/2008/37?limit=1>):

```
<as-nodes table-name="ASNodesTbl_2008_week_37" total-nodes-count="19002" offset="0"
limit="1">
<as-node as-number="6067" agent-who-discovered="16166" date-of-discovery="2008-09-08
11:40:30.0" date-of-validation="2008-09-08 11:40:30.0"/>
</as-nodes>
```

## 3. AS Edges Data:

1. Represents edges (links) between nodes in different Autonomous Systems (Ases).
2. Each row (AS Edge) includes this data:
  1. id - running id of the row.
  2. source-as - The AS number of the source of this edge.
  3. dest-as - The AS number of the end point of this edge.
  4. agent-who-discovered – see [section 2.2.2](#).
  5. date-of-discovery – see [section 2.2.3](#).
  6. date-of-validation – see [section 2.2.4](#).
  7. visit-count – number of times this edge was seen in the topology this week.
  8. validating-agent – index of the last DIMES agent who saw this edge in this week.

### 3. Example of data

(URL: <http://www.netdimes.org/rest/topologies/asedges/2008/37?limit=1>):

```
<as-edges table-name="ASEdgesTbl_2008_week_37" total-edges-count="48720" offset="0"
limit="1">
<as-edge id="1" source-as="30880" dest-as="12381" agent-who-discovered="15277" date-of-
```



```
discovery="2008-09-08 13:26:29.0" date-of-validation="2008-09-14 16:12:54.0" visit-count="6"
validating-agent="14070"/>
</as-edges>
```

#### 4. Nodes Data:

1. Represents nodes in IP level.
2. Each row (IP node) includes this data:
  1. id – see [section 3.2.1](#).
  2. ip – IP address of the node. It can happen that a node has an unknown IP (anonymous node). In case like that the address will appear as IP1-X?Y-IP2 where X stands for the distance from IP1 and Y stands for the distance from IP2. For example, you can see an address like: IP1-10?3-IP2, which means that this is the 10th unknown router and there are 2 more after it (3 including it). IP1 represents the last known address before and IP2 represents the first known address after it.
  3. host-name – host name of this node. If the name is unknown the IP number will appear or null.
  4. agent-who-discovered – see [section 2.2.2](#).
  5. as-prefix – the AS prefix to which this node belongs to.
  6. date-of-discovery – see [section 2.2.3](#).
  7. date-of-validation – see [section 2.2.4](#).
  8. as-number – see [section 2.2.1](#).
3. Example

(URL: <http://www.netdimes.org/rest/topologies/nodes/2008/37?offset=5&limit=1>):

```
<nodes table-name="NodesTbl_2008_week_37" total-nodes-count="540254" offset="5" limit="1">
<node id="6" ip="195.50.113.18-1?2-83.229.32.253" host-name="null"
agent-who-discovered="12257" as-prefix="Unknown" date-of-discovery="2008-09-08 12:03:03.0"
date-of-validation="2008-09-08 12:03:03.0" as-number="-1"/>
</nodes>
```

#### 5. Edges Data:

1. Represents edges (links) between IP nodes.
2. Each row (IP level edge) includes this data:
  1. id - see [section 3.2.1](#).
  2. source-ip – the source IP of the edge (beginning of edge). See also [section 4.2.2](#) for



details about the structure of the field.

3. destination-ip – the destination IP of the edge (end of edge). See also [section 4.2.2](#) for details about the structure of the field.
4. date-of-discovery – see [section 2.2.3](#).
5. date-of-validation – see [section 2.2.4](#).
6. agent-who-discovered – see [section 2.2.2](#).
7. inter-as – either this an edge between two IPs in the same AS (this case the value will be 0), between 2 different Ases (value will be 1) or is it unknown (value will be -1).
8. is-unknown – A flag to report wether a router answered with a valid IP address. Unknown address will appear with the anonymous node pattern (see [section 4.2.2](#)).

Options of the field are:

1. 0 = both source-ip and destination-ip have a valid IP address.
2. 1 = source-ip is anonymous node.
3. 2 = destination-ip is anonymous node.
4. 3 = both source-ip and destination-ip are anonymous nodes.
9. min-delay – since each masurement in the DIMES Agents runs more than once here you can see the minimum delay – out of the few runs - between the 2 nodes in the edge.
10. average-delay – the average delay between the 2 nodes in the edge, out of the runs of the measurement (note that due to traceroute pathologies this can be negative).
11. measurements-count – number of time this edge was seen during the week of the topology.

### 3. Example of data

(URL: <http://www.netdimes.org/rest/topologies/edges/2008/52?limit=1>):

```
<edges table-name="EdgesTbl_2008_week_52" total-edges-count="1601314" offset="0"
limit="1">
  <edge id="1" source-ip="80.236.7.118" destination-ip="81.65.0.1"
date-of-discovery="2008-12-23 11:24:43.0" date-of-validation="2008-12-31 14:14:19.0"
agent-who-discovered="18233" inter-AS="1" is-unknown="0" min-delay="-3"
average-delay="-4.0" measurements-count="15"/>
</edges>
```

### 6. City Nodes Data:

1. Represents data about cities which was collected by DIMES Agents.



2. Each row (City Node) includes this data:
  1. id - see [section 3.2.1](#).
  2. As-set – Numbers of all the ASes which can be found in this city.
  3. city – City name.
  4. country – The country to which this city belongses.
  5. date-of-discovery – see [section 2.2.3](#).
  6. date-of-validation – see [section 2.2.4](#).
  7. latitude – The geographical latitude of the city.
  8. longitude – The geographical longitude of the city.
  9. num-of-ips – Number of IPs which where found in this city during DIMES measurements.

3. Example of data

(URL: <http://www.netdimes.org/rest/topologies/citynodes/2009/2?limit=1>)

```
<city-nodes table-name="CityNodesTbl_2009_week_2" total-nodes-count="17761" offset="0"
limit="1">
<city-node id="1" AS-
Set="8801,7029,21217,3356,28959,8391,8327,41623,6853,513,1297,6837,43189,4323,12333,3303,8659,
559,35071,10764,15361,8928" city="Geneva" country="Switzerland"
Date-of-Discovery="2009-01-05 12:01:12.0" Date-of-Validation="2009-01-11 10:59:28.0"
latitude="46.2" longitude="6.1667" num-of-ips="87"/>
</City-Nodes>
```

7. City Edges Data:

1. Represents connections (links) between different cities which were identified by DIMES Agents.
2. Each row (City Edge) includes this data:
  1. id - see [section 3.2.1](#).
  2. as-edges-set – In case that source and destination nodes of the edge are not part of the same AS the relevant ASes will appear in pattern of Source AS – Destination AS.
  3. as-set – In case that both source and destination nodes of the edge belong to the same AS its number will appear here.
  4. destination-city – Name of the end point city.
  5. destination-country – Name of the country which the end point city belongs to.



6. source-city – Name of the start point city.
  7. source-country – Name of the country which the start point city belongs to.
  8. date-of-discovery - see [section 2.2.3](#).
  9. date-of-validation - see [section 2.2.4](#).
  10. num-of-ip-edges – Number of IP edges between those cities.
3. Example of data (URL: <http://www.netdimes.org/rest/topologies/cityedges/2009/9?limit=1>)

```
<city-edges table-name="CityEdgesTbl_2009_week_9" total-edges-count="88937" offset="0"
limit="1">
<city-edge id="1" as-edges-set="26554-7949" as-set="" destination-city="Culver" destination-
country="United States" source-city="Ann Arbor" source-country="United States" date-of-
discovery="2009-02-23 12:09:09.0" date-of-validation="2009-03-01 08:25:13.0" num-of-ip-
edges="1"/></city-edges>
```

8. Routers data:

1. Represents data about routers which was collected in DIMES Ping Measurements.
2. Each row (Router) includes this data:
  1. id – Router id in table. In case that a router has more than one interface (e.g IP number) the same id will appear more than once. For example: If router number 2356 has 3 interfaces there will be 3 rows with the same id – 2356 – each row with a different IP number.
  2. ip – the IP of this interface.
  3. size – in case that this router has more than one interface this number will be bigger than one. The following rows will have the other interfaces with the same id.
3. Data example (URL: <http://www.netdimes.org/rest/topologies/routers/2009/9?limit=1>)

```
<routers table-name="RoutersTbl_2009_week_9" total-routers-count="1079657" offset="0"
limit="1">
<router id="284586" ip="207.89.129.94" size="1"/>
</routers>
```