1 Scope

This document defines the protocol for accessing resources on a remote system.

2 Normative references

IANA Media Type Reference
<http://www.iana.org/assignments/media-types/media-types.xhtml>
IETF RFC 4122 - A Universally Unique IDentifier (UUID) URN Namespace
ONVIF Core Specification

3 Terms and Definitions

3.1 Definitions

Media Configuration
An abstract component that produces or consumes a media stream on the network, i.e. video and/or audio stream.

Media Profile
Maps video and audio sources and outputs encoders as well as PTZ and analytics configurations.

Remote Token
Token of a remote resource

Resource
In this document the term is used to refer to an ONVIF resource that can be addressed via a token.

Token
Unique textual reference of a resource.

3.2 Abbreviations

VMS
Video Management System (synonym for VSS)

VSS
Video Surveillance System

4 Overview

This document describes how a Video Management System can expose its resources so that a remote VSS or a client can make use of them with the APIs defined by the various ONVIF specifications.

The example in Figure 1 shows two cameras attached to a lower level VSS which are operated from a client that connects to an upper level VSS. This document assumes that standard existing ONVIF commands are used for e.g. moving the focus in the example below.

This specification focuses on resource naming and querying.
4.1 Resource Addressing

In the ONVIF specifications resources are addressed by so called tokens. Tokens are character strings of a defined length and are enumerated by the device to ease the devices resource management. Similarly this specifications assumes that a VSS enumerates its attached device resources in such a way that a unique token is assigned to each of its attached devices resources.

Additionally this specification assumes that a VSS implements resource token mapping by adding a prefix delimited by a colon. Whether a VSS simply prepends a prefix to a device token or does a complete remapping is outside of the scope of this specification.

The following list provides an incomplete list of resources that can be handled:

- Media Profiles
- Video and Audio Sources
- Any Media Configurations including OSD and Masks
- PTZ Nodes
- Digital Inputs
- Relay Outputs
- Recordings, Recording Tracks and Recording Jobs

4.2 Resource Queries and Updates

This specification uses the ONVIF property event mechanism to model global resources. A client or upper level VSS subscribes to a VSS pull point. In consecutive requests it then pulls all resources it is interested in in order to get to know all relevant resources of a lower level VSS.

Once all resources have been reported further pull messages will either timeout when no changes happened or report changes. In this context changes are added resources, modification of resource properties or removed resources.

By applying the property event notification mechanism to resources a client or upper level VSS has always up to date information about the lower level VSS resources it is interested in.
4.3 Events
An upper level VSS can retrieve events either generic or for a given token. In order to retrieve e.g. motion events from a dedicated resource a client has to provide a content filter with matching source token.

4.4 Authentication
Access to remote VSS may require strict authentication constraints. Refer to the ONVIF Security Configuration Specification for means to server and client authentication.

4.5 Live Checks
The mechanism defined in this specification make use of the ONVIF real-time pull point. When no events are pending at a pull point the PullMessages call will timeout after a client defined timeout. The client is in control of the timeout and can so ensure a continuous heartbeat from the server. Hence an additional ping or heartbeat interface is not needed.

4.6 Live Video Example
This example assumes that a police station client wants to get live feeds related to an incident at a certain geo location.

1. The client authenticates itself at the VSS using its client certificate.
2. Client subscribes to VSS topic Resource/MediaProfile in a distance of 100 meter of a lon/lat provided geo location.
3. VSS responds to PullMessages with enumerating all media profiles in the given area.
4. If the client receives multiple profiles for a remote device it may select the best suited one by querying the profile configurations. For each profile token the client calls GetProfiles. By supplying parameter Type set to VideoSource and VideoEncoder it will receive detailed information about the Video resolution and encoder settings.
5. The client calls GetStreamUri for each video source on the selected profile
6. The client streams video and displays it on the Police Video Wall.

At the first glance it looks like no new functional interface has been defined. Although this is correct, the VSS still has to implement a number of operations in order to make the above example happen:

a. Enumerate its attached camera profiles as resource events
b. Implement the Media2 API
c. Map request media profile token to cameras and their local tokens
d. Map response profile tokens to global tokens
e. Support streaming proxy and map camera stream uri to proxy

Additionally an intelligent VSS may expose only those media profiles to its client that have a decent data rate so that it is able to proxy those Video streams.

Instead of directly connecting to the remote lower level VSS systems the police station client may connect to an upper level VSS that in turn combines search results from multiple lower level VSS and forwards commands to individual camera sources via the appropriate lower level VSS.

4.7 Forensic Example
This example assumes that a police station client wants to retrieve recordings related to an incident at a certain geo location.

1. The client authenticates itself at the VSS using its client certificate.
2. Client subscribes to VSS topic Resource/VideoSource in a distance of 100 meter of a lon/lat provided geo location.

3. VSS responds to PullMessages with enumerating all video sources in the given area.

4. Client calls FindEvents passing the interested time range and a list of video sources as search scope.

5. Client calls GetEventSearchResults to get all available recordings at the interested location and in the given period of time.

6. The client retrieves the RTSP URIs by calling the Replay GetReplayUri method for the recording.

7. The client plays back video of interested cameras by retrieving the RTSP media stream for the uri retrieved by a call to Replay:GetReplayUri.

Alternative to the above approach the client might also directly query recording resource at a geo location. Some replays may fail because the cameras did not record for the selected location at the selected time of day.

Above example is applicable for both content stored in camera's edge storage and VMS recording server. It is up to the VMS application how the recordings are retrieved.

4.8 Notification Mechanism

This specification assumes that the ONVIF Realtime Pullpoint mechanism is deployed. Note that the resource query is transparent to the event transport mechanism and the events may also be conveyed via the OASIS Base Notification mechanism or the Notification Streaming Interface.
5 Resource Addressing

5.1 Remote Tokens

ONVIF specifications are generally assuming that the device defines tokens which are unique within a device and its context. This specification extends the scheme to allow building globally unique tokens called remote tokens.

A remote token shall be constructed like a qname with a device specific prefix and a local token.

RemoteToken = Prefix + ':' + LocalToken

The overall string length of the remote token is limited to 64 characters. A local token must not exceed 36 characters and should contain no colon. The length limitation is chosen such that it enables the use of UUIDs as defined in RFC 4122. Note that device implementations typically use compact tokens composed from few characters.

A VSS shall use the same prefix for all tokens of the same device. This allows a client to understand which tokens they can use for any ONVIF API call.

A VSS may choose to simply use device local tokens as LocalToken part or create an internal mapping. A client may not assume that use tokens received from a VSS directly in device calls by stripping the prefix.

The naming conventions for the prefix part are outside of the scope of this specification. Depending on the application area implementers may choose different approaches. As a consequence this specification does not mandate that remote tokens are globally unique between different VSS. See Annex A for a country specific definition of globally unique addresses.

5.2 Token Context

Clients talking to multiple servers at a time like VSS and/or devices shall address resources to a server only with tokens received from that same server. There is no guarantee that remote tokens received from one server may be used to address the same resource at another server.
6 Resource Queries

6.1 General

This specification models resources as so called property events. A resource is an ONVIF configuration item addressed via a token.

A server supporting resource queries shall signal the supported resource queries via the GetEventProperties interface of the event service. The following resources may be enumerated:

For media configurations:
- VideoSourceConfiguration
- AudioSourceConfiguration
- VideoEncoderConfiguration
- AudioEncoderConfiguration
- AudioOutputConfiguration
- AudioDecoderConfiguration
- MetadataConfiguration
- AnalyticsConfiguration
- PTZConfiguration
- OSDConfiguration
- MaskConfiguration

For media profiles: MediaProfile

6.2 Resource Event

Each resource maps to the following event definition;

```xml
<T:MessageDescription IsProperty="true">
  <T:Source>
    <T:SimpleItemDescription Name="Token" Type="T:ReferenceToken"/>
  </T:Source>
  <T:Data>
    <T:SimpleItemDescription Name="Name" Type="xs:string"/>
    <T:ElementItemDescription Name="Location" Type="T:GeoLocation"/>
    <T:SimpleItemDescription Name="Scope" Type="xs:string"/>
    <T:SimpleItemDescription Name="Offline" Type="xs:boolean"/>
  </T:Data>
</T:MessageDescription>
```

The source item Token is mandatory and shall contain a qualified token that is unique within the serving system.

The data items are optional.

The data item scope refers to the ONVIF device discovery scope entry. It may occur multiple times for each scope entry supported by the device.

An event shall be generated with PropertyOperation set to Initialized whenever a resource is signaled the first time in a subscription or it is newly added to the system. An event with PropertyOperation set to Deleted shall be generated when a resource is removed from the system.

Note that a change from online to offline or vice versa shall only create a PropertyOperation of type Changed if the event contains an Offline state boolean.

6.3 Location Filter

A service supporting resource queries shall support the Location Filter.

```xml
<xs:complexType name="LocationFilter">
  <xs:sequence>
    <xs:element name="lon" type="xs:double"/>
    <xs:element name="lat" type="xs:double"/>
    <xs:element name="height" type="xs:float" minOccurs="0"/>
    <xs:element name="radius" type="xs:float"/>
    <xs:element name="includeUnknown" type="xs:boolean" minOccurs="0"/>
  </xs:sequence>
</xs:complexType>
```
Devices with unknown location shall correspond to a match if the property includeUnknown is set.

6.4 Prefix Filter

A server supporting resource queries shall support the Prefix Filter. The prefix filter allows to restrict the search to any events of a device by the prefix assigned by the VMS.

```xml
<xs:complexType name="PrefixFilter">
  <xs:sequence>
    <xs:element name="Prefix" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
```

6.5 Scope Filter

A server supporting resource queries shall support the Scope Filter allowing to search for device scope entries.

```xml
<xs:complexType name="ScopeFilter">
  <xs:sequence>
    <xs:element name="Scope" maxOccurs="unbounded">
      <xs:complexType>
        <xs:simpleContent>
          <xs:extension base="xs:string">
            <xs:attribute name="Match"/>
          </xs:extension>
        </xs:simpleContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
```

See section 7.3.3 of the ONVIF Core Specification for the scope matching rules.

6.6 Select Filter

A server supporting resource queries shall support the Select Filter allowing to restrict the resulting data items of each event in order to reduce the message sizes of large queries.

```xml
<xs:complexType name="SelectFilter">
  <xs:sequence>
    <xs:element name="Path" type="xs:string" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>
```

Each entry defines an XPath expression that matches to one of the event data items.

The following example restricts the content of the result to the resource name and location scope:

```xml
<SelectFilter>
  <Path>/Name</Path>
  <Path>/Scope/Location</Path>
</SelectFilter>
```
Annex A. Addressing Scheme
(informative)

This Annex defines a globally unique addressing scheme that can be used to address devices like cameras, DVRs and others. The scheme includes information for area, device type.

A.1 Overview

Figure A1 shows the order of the four fields in the 20-digit encoding scheme:

![Figure A1: Fields of the device ID scheme](image)

The device ID and ID of other resources should finish the initial configuration in the bottom platform.

A.2 Field Definitions

A.2.1 Zone Code

The eight digit zone code is split into four two digit items as shown in Table A1.

<table>
<thead>
<tr>
<th>2 digit</th>
<th>2 digit</th>
<th>2 digit</th>
<th>2 digit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Province</td>
<td>City</td>
<td>District</td>
<td>Unit</td>
</tr>
</tbody>
</table>

The first six digit is the Administrative Division code, which depends on the location of the Surveillance Center and meets the requirement of GB/T 2260-2007, similar to the post code. For example, the code of city X is 420100.

The unit code is defined by users, such as Police Station etc. It starts from 01, 00 represents the Surveillance Center.

A.2.2 Agency Code

Table A2 defines the two digit agency codes.

<table>
<thead>
<tr>
<th>Code Type</th>
<th>Name</th>
<th>Subject</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Urban Roadway</td>
<td></td>
<td>includes urban roadway, commercial avenue, public areas &amp; key areas, etc.</td>
</tr>
<tr>
<td>01</td>
<td>Society Community</td>
<td></td>
<td>includes residential districts, buildings, and cyber café, etc.</td>
</tr>
<tr>
<td>02</td>
<td>Social Security Agency</td>
<td></td>
<td>includes public security buildings &amp; detention room</td>
</tr>
<tr>
<td>03</td>
<td>Other Security Agencies</td>
<td>Government Offices</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>Traffic Roadway</td>
<td></td>
<td>includes road artery, national roadway &amp; highway</td>
</tr>
<tr>
<td>05</td>
<td>Traffic Checkpoints</td>
<td></td>
<td>includes crossroads, E-police, checkpoint &amp; toll gate, etc.</td>
</tr>
<tr>
<td>Code</td>
<td>Device Type</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------------------</td>
<td>--------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>DVR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>Video Server</td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>Encoder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>Decoder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>Video Switching Matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>Audio Switching Matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>Alarm Controller</td>
<td></td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>NVR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>Online Information Acquisition Device</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>Online Information Acquisition System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>Checkpoints on Roadway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>HVR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>121-129</td>
<td>Extended Front-end devices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>Camera</td>
<td></td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>IPC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A.2.3 **Device Type Code**

Table A3 defines the three digit device type codes.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>133</td>
<td>Display Encoding</td>
</tr>
</tbody>
</table>
| 134 | Alarm Input Devices  
     | (IR alarm, smoke sensor alarm & access control, etc.) |
| 135 | Alarm Output Devices  
     | (Alarm bell & Alarm lamp ) |
| 136 | Audio Input Devices |
| 137 | Audio Output Devices |
| 138 | Mobile Transmission Devices |
| 139 | Other Peripheral Devices |
| 140-149 | Types of Extended Peripheral Devices |
| 200 | Central Signaling Control Server  
     | (Signal Networking Unit)  
     | 200-299 indicates platform devices. |
| 201 | Web Application Server |
| 202 | Media Distribution Server |
| 203 | Proxy Server |
| 204 | Security Server |
| 205 | Alarm Server |
| 206 | Database Server |
| 207 | GIS Server |
| 208 | Management Server |
| 209 | Access Gateway |
| 210 | Media Storage Server |
| 211 |   |
| 215 | Traffic Grouping |
| 216 | Virtual Grouping |
| 212-213 |   |
| 217-299 | Extended Device Type |
| 300 | Central User  
     | 300-399 indicates the central user. |
| 301-343 | Industrial Role User |
| 344-399 | Extended Central User Type |
| 400 | Terminal User  
     | 400-499 indicates terminal user. |
| 401-443 | Industrial Role User |
| 444-499 | Extended Central User Type |
| 500 | Signaling Server of Video/Image Information Application Platform  
     | 500-599 indicates the external servers of platform. |
| 501 | Signaling Server of Video/Image Information Maintenance Platform |
| 502 | Video/Image Analysis Device/System of Public Security  
     | Newly-added in Video/Image Library |
| 503 | Video/Image Information Database of Public Security  
     | Newly-added in Video/Image Library |
| 504 | Video/Image Information Platform of Public Security  
     | Newly-added in Video/Image Library |
| 505-599 | Extended Platform Server (External) |
| 600-699 | Extended Type  
     | 600-699 indicates the extended types. |
A.2.4 **Serial Number**

The seven digit serial number generally follows numerical order, which includes the location information, such as Subway-Subway Passage-Ticket Office-Platform-Subway Carriage.

<table>
<thead>
<tr>
<th>1 digit</th>
<th>6 digit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network ID</strong></td>
<td><strong>Serial number of device or user</strong></td>
</tr>
</tbody>
</table>

The Network ID is encoded as follows:
- 0, 1, 2, 3 and 4 refer to surveillance alarm network;
- 5 refers to public security network;
- 6 refers to government network;
- 7 refers to Internet;
- 8 & 9 are reserved;

A.2.5 **Examples**

IP Camera No.15 of Metro Line 1’s No.10 ticket-check in city X:

<table>
<thead>
<tr>
<th>City X</th>
<th>Metro Line 1</th>
<th>Agency Code</th>
<th>Device Type Code</th>
<th>Station No.</th>
<th>Ticket Check</th>
<th>IPC No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>320100</td>
<td>01</td>
<td>07</td>
<td>132</td>
<td>10</td>
<td>08</td>
<td>015</td>
</tr>
</tbody>
</table>

Server No.02 of Metro Line 1’s surveillance center in city X:

<table>
<thead>
<tr>
<th>City X</th>
<th>Line 1</th>
<th>Agency Code</th>
<th>Device Type Code</th>
<th>Invalid Station No.</th>
<th>Invalid Ticket Check</th>
<th>Server No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>320100</td>
<td>01</td>
<td>07</td>
<td>200</td>
<td>00</td>
<td>00</td>
<td>002</td>
</tr>
</tbody>
</table>
Annex B. APIs with Token Adaption

(Informative)

This annex informs about the expected resource token handling of a VSS. Incoming global token may need to be transformed into device local tokens and tokens received from a device may need to be transformed into global tokens.

B.1 Consuming Live Video

- Media2:GetProfiles
  Both incoming and outgoing profile and configuration tokens need to be adapted.

- Media2:GetStreamingUri
  The incoming profile token needs to be adapted. Additionally the VSS needs to ensure that the streaming URI can be reached by the client. This shall include a streaming proxy service.

- Media2:GetSnapshotUri
  The incoming profile token needs to be adapted. Additionally the VSS needs to ensure that the snapshot URI can be reached by the client. This shall include a proxy service.

B.2 Controlling PTZ Cameras

For all the below listed functions the incoming profile token needs to be adapted.

- PTZ:AbsoluteMove, PTZ:ContinuousMove
- PTZ:Stop
- PTZ:GetPresets, PTZ:GetPreset
- PTZ:SendAuxiliaryCommand

B.3 Retrieving Recordings

- RecordingSearch:GetRecordingInformation
  Both incoming and outgoing recording and track tokens need to be adapted.

- RecordingSearch:FindRecordings
  Both incoming and outgoing recording, track and search tokens need to be adapted.

- RecordingSearch:GetRecordingSearchResults
  Both incoming and outgoing recording, track and search tokens need to be adapted.

- RecordingSearch:EndSearch
  The incoming search tokens needs to be adapted.

- Replay:GetReplayUri
  The incoming recording token needs to be adapted. Additionally the VSS needs to ensure that the streaming URI can be reached by the client. This shall include a streaming proxy service.

B.4 Forwarding of Events

Typically the Source section contains tokens that need to be adapted.
## Annex B. Revision History

<table>
<thead>
<tr>
<th>Rev.</th>
<th>Date</th>
<th>Editor</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.12</td>
<td>Dec 2018</td>
<td>Hans Busch</td>
<td>First release</td>
</tr>
</tbody>
</table>
Bibliography

ONVIF PTZ Service Specification
ONVIF Security Configuration Specification
ONVIF Media2 Service Specification
ONVIF Recording Search Service Specification
ONVIF Replay Control Service Specification
ONVIF Streaming Specification