PCP Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 08, 2014

H. Moustafa
D. Moses
Intel Corporation
M. Boucadair
France Telecom
November 04, 2013

PCP Extension for Signaling Feedback Information from the End-User Application to the Application Sever and to the Network
draft-mou-pcp-application-network-feedback-00

Abstract

Nowadays users consumption style for video and multimedia applications is strongly changing. Users are heavily counting on wireless and mobile devices for video streaming and interactive video and multimedia applications. This can be implemented for instance by having more intelligence in the service and the network infrastructure to better suit a differentiated users consumption style. This can be achieved through having: (i) a knowledge in the network and service platform about the available device and network conditions for the end-user and (ii) a knowledge in the network about the content requirements in terms of devices capabilities and network resources for content stored either in the network or in the application server. To obtain such knowledge with no need for changing the current infrastructure and in a generalized way to all applications, feedback/notification mechanisms between the end-user application, the network and the service platform is needed to provide information helping the content delivery and adaptation decisions. This document is investigating such application-agnostic track.

This document extends the Port Control Protocol (PCP) RFC 6887 [RFC6887] to allow: (i) the users application to notify the network and application server about its available resources in terms of devices capabilities and network conditions as well as information about the user (e.g., location, mobility status) and (ii) the application server to notify the network about the requirements of the content it stores in terms of devices and network resources. A new PCP option, denoted the FEEDBACK, is specified to allow such feedback notification signaling. This option is used with both PEER and MAP Opcodes.

Status of This Memo

This Internet-Draft will expire on May 08, 2014.

Copyright Notice

Table of Contents

1. Introduction

Nowadays, Internet traffic includes huge amount of video traffic which is expected to dominate the worlds mobile traffic in the coming years (66% of mobile video traffic according to Cisco forecast). The users consumption style for video services is strongly evolving towards a user-centric model enabling video services access based on users profiles and making receiving device (e.g., Laptops, tablets, smartphones, and future devices models) constitute the majority of end-user devices for video services through plenty of services over the Internet.

Although this big evolution, the network and service infrastructure are not optimized enough to handle the content delivery and adaptation function of the network resources, devices resources, application and content requirements. As a result, Quality of Experience (QoE) for video services and multimedia applications ca not be guaranteed in some situations.

This document defines an extension to the Port Control Protocol (PCP) RFC 6887 [RFC6887] allowing: (i) the end-user application to signal in real-time to the network and application server information about its available device capabilities and network resources (mainly device characteristics, buffering status as an indication of the network conditions as well as other useful context information (e.g., location, environment light/noise, mobility status) and (ii) the application server to signal in real-time to the network the requirement of the content it stores in terms of devices and network resources. The extension defines a new PCP option for the existing PEER and MAP OpCodes: FEEDBACK Option for signaling information between the end-users application, the network and the application server.

Motivations to undertake this effort are discussed in Section 3 through a number of use cases, while justifications for the use of PCP are elaborated in Section 4.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Use Cases

The new model for video services consumption over mobile and wireless networks creates the need for feedback information between the end-user application, the network and the service platform in real-time. This helps optimizing the content for device/network resources on one hand and end-user QoE and battery experience on the other hand. In this view more intelligence needs to be considered in the network. A way to do that is through having several service functions in the network middle boxes on the path between the end-user and the application server, as shown in Draft [ServiceFunctions], that could be collocated with the NAT function or not. Examples of service functions that we consider to better manage the video delivery are: video optimization function (including transcoding), caching capabilities, TCP optimizer, video controller allowing for example content recommendation or intelligent scheduling in the case of networks. The following subsections present some related use-cases:

3.1. Optimized Content Delivery by the Network

In this case, the end-users device is the host implementing the PCP client and a middle box network node in the network (e.g. edge router, home gateway or any cache node close to the user) is the PCP server. This middle box node can store a copy of the content or can be a content relay from the application server to the end-user. If the middle box network node is only a relay, then it receives the feedback information sent to the network by the end-user application and by the application server and makes use of this information for optimizing the content it relays to the end-user. If the middle box network node stores the content, then it makes use of the feedback information sent to the network by the end-user application to optimize the stored content when requested by the end-user. This use-case is beneficial in adaptive video streaming and server-based video conferencing applications. In the former, the middle box network node can be an edge router, home gateway or any network node caching the content. In the latter, the Multi-Point Control Unit (MCU), with which video conferencing clients communicate, is the middle box network node.

3.2. Optimized Content Delivery by the Service Platform

In this case, the end-user application is the host implementing the PCP client and the application server implements a PCP server. Feedback information is sent from the end-user application to the application server allowing the server to make intelligence decisions on the content adaptation. This use-case applies also in a Content Delivery Network (CDN) case, where several application servers exist and an application server controller, who plays the role of a PCP proxy, directs the end-user request to the appropriate application server (function of proximity for instance or load balancing).

3.3. Network-based Video Session Seamless Experience across Devices

Device awareness in terms of the device capabilities enables application content to be transferred from device to device or to be shared across multiple devices. In this case, the devices are located behind the same residential gateway, and therefore be reachable from outside with the same IPv4 address or IPv6 prefix. Each end-user device is a host implementing PCP client and sends feedback information on its devices resources to the middle box network node (e.g. CPE or home gateway) playing the role of PCP server. The functionalities related to the session transfer from device to device or sharing the same content across multiple devices is outside the scope of this document. Our focus is how to benefit from the feedback information during these cases as shown below:

If a video session started on device1 (e.g. smartphone) and is transferred to device2 (e.g. a tablet or TV screen) when it becomes available in the proximity of device1, the end-user application feedback information sent by device 2 will be used by the middle box network node to adapt the application content to match the device resources and network conditions of device 2.

If the same video is watched by multiple users with different devices resources and network conditions (e.g., video for a lecture in a classroom), the end-user application feedback information sent by each user device will be used by the middle box network node to adapt the application content to match the device resources and network conditions for each user device.

3.4. Network-based User-Centric Video Content Adaptation

In this case, the end-user device is the host implementing the PCP client, the application server storing the content is a PCP client , and a middle box network node that stores and/or relays the content implements a PCP server and receives feedback from both the end-user application and the application server. The feedback information (mainly on the location, battery level and mobility status) sent by the end-user application to the middle box network node is used in the decision of content adaptation (mainly the insertion of targeted advertisements). If the middle box network node does not store the content, then it can receive the native content from the application server then adapts it making use of the feedback information. If the middle box network node stores the content then it adapts it directly for each user.

3.5. Optimization Examples

The following are some examples on how to optimize the content for device and network resources on one hand and end-user QoE and battery experience on the other hand.

- For a high display resolution, a video of high bitrate/resolution is sent if the users application buffering level and the global network bandwidth conditions are sufficient and the battery level is sufficient.
- If the battery level degrades during the session even if the users application buffering level remains sufficient, then the video is continued to be streamed in lesser bitrate/resolution to save battery and prevent video session interruption due to battery failure (keeping a threshold on the quality based on the remaining resources). In this case, lesser resources can be allocated for cellular users to match the bandwidth required for the low bitrate video.
- If the battery level is fine but the users application buffering level and global network bandwidth conditions degrades, the video switches to lower bitrate (following the ordinary adaptive streaming approach), which in turns save in the battery level.
- For small size display, the transmission errors are less observed by the user especially if the user is mobile and is in a noisy environment and so video can be displayed with lesser quality (bitrate) to save battery resources and network resources even if the buffering level and global network bandwidth conditions are not poor. This is much useful for content categories as news and non-live content, in which transmission errors have lesser impact on the users perception.
- If the user is in a bright/sunny environment and the battery level is not high, content can be sent with higher brightness but with lower bitrate to compensate the power consumed from content brightness.
- If the user is in a dimly environment, no matter of the battery level and the global network bandwidth, the user can receive content with lower brightness. This has a dual benefit: allowing better perceived quality for the user and saving in battery resources.
- Users can receive content including targeted advertisement for regional services according to each user location, mobility status and battery level suitability. This saves network bandwidth and devices resources (mainly battery resources) happening from the reception of bigger traffic volume if each user receives all the advertisements.

4. Why PCP?

The use of PCP to signal feedback/notification between the end-users application, the network and service platform in real-time is motivated by the following: In the Port Control Protocol (PCP) specification RFC 6887 [RFC6887], PCP is viewed as a request/response protocol and also as a hint/notification protocol between a PCP client and a PCP server. Message flows in PCP are viewed as independent streams carrying information between the PCP client and the PCP server, in which the PCP client sends a stream of hints indicating to its server its mapping status and the PCP server sends to the client a notification informing the client on the actual state of its mapping. In this view of the protocol, PCP can be extended to carry more mapping information than the IP internal versus external addresses. Draft [Flowdata] is an example of the use of PCP for signaling by the client the flow characteristics to the network and signaling by the network its ability to accommodate that flow back to the client.

In addition, PCP allows learning and influencing the mapping lifetime, which helps reducing network bandwidth, overload on application servers and middle boxes and battery resources for wireless and mobile devices. This makes PCP suitable for conveying feedback information in our use-cases in real-time and resource wise way.

Further advantages for PCP motivating its use are as follows:

- PCP can be used to install state in upstream devices such as NAT, firewalls or other flow-aware devices.
- PCP can be used to notify a failure that may occur at an upstream PCP-controlled device, and therefore the PCP client can react accordingly.
- PCP allows learning the lifetime of installed mappings and would therefore avoid overloading the network and service platform with keepalive messages. This also saves the battery resources for wireless and mobile end-user devices.
- PCP can be used to notify the network with the flow characteristics so as to enforce policies at the access segment.
- PCP can be used to receive informative information from the network so that client may use them to select the interface to use to place a session.
- PCP can be extended easily to allow reporting capabilities to a remote server.
- Extending PCP with the FEEDBACK feature avoids making assumptions on how media streams are exchanged (e.g., RTP, IAX mini-frames, etc.).
- PCP extension does not require an OS support. The feature can be managed at the application level.

5. FEEDBACK PCP Option

This document presents an extension to PCP through the definition of FEEDBACK option in PCP. The FEEDBACK option aims to signal feedback information between the end-user application, the network and service platform to help optimizing the network and devices resources and enhancing the users experience. This feedback mechanism makes also use of the PCP FLOWDATA option [Flowdata]. This means that the PCP client sends both FLOWDATA and the FEEDBACK options in the same requests.

The information signaled in the FEEDBACK Option may include the screen size, screen resolution and battery level as device capabilities information and the users location, environment type and mobility status as context information on the user side. This information is signaled once at the beginning of the session and can be updated upon variation. Following the information signaling, the PCP server uses the FEEDBACK option to signal its ability of providing content with characteristics matching the device and network resources as well as the user context. This information will be used by the application server and by the network (through the middle box network nodes having service functions) for optimized content adaptation as explained in Section 3.

The FEEDBACK option does not make any assumption on the presence of NAT and/or firewall. In particular, PCP-based mechanism to instantiate state in an upstream NAT, firewall and any other flow-aware device are not impacted by the use of FEEDBACK option.

The PCP client is implemented at the end-user device and the application server (content server), and the PCP server is implemented at the middle box network node storing the content and the application server. The PCP client may be configured with multiple IP addresses; each of them pointing to a distinct PCP server. The PCP client will contact all these PCP server in parallel as discussed in [I-D.ietf-pcp-server-selection].

A PCP Proxy [PCPProxy] functionality can be enabled in intermediate nodes (e.g., Customer Premise Equipment (CPE)) on the path between the PCP client and the PCP server.

Upon receipt of the FEEDBACK option by the PCP Server, its content is passed to the Application Server that decides whether an action is needed to serve the requesting client to better accommodate its resources and conditions.

The procedure for generating a request that includes the FEEDBACK option, handling a request that includes the FEEDBACK option, and generating a response to a request that includes the FEEDBACK option is similar to the behavior specified in [Flowdata].

Triggers to generate a request that capture the network conditions and device recourses in a FLOWDATA and FEEDBACK options are local to the application. How the application server or network middle boxes makes use of the content of the FLOWDATA and FEEDBACK options is also local to the PCP Server and to each the decision-making process at the Application Server side or network middle box node.

5.1. PCP Request and Responses

The PCP client follows the steps of generating the PEER and MAP opcodes request and response as described in [Flowdata]. The FEEDBACK option is included in the request and response according to the format described in this section.

Option Name: FEEDBACK
Number: to be assigned by IANA
Purpose: Describe to the network information on the end-user application and user context (e.g., device characteristics, buffering status as indication of the network conditions, location, environment light/noise, mobility status), and information on the content that can be provided by the application server.
Valid for opcodes: PEER and MAP
Length: 16 Octets
May appear in: request. May appear in response only if it appeared in the associated request
Maximum occurrences: 1

The FEEDBACK option request has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Option code=TBA|Reserved Option|    Option Length= 16          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Reserved                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                SRPW                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                SRPH                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   BLS |   SSz |EnvNs|EnvLt|MStat|      Location               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: FEEDBACK Option Request

The fields are described as follows:

- SRPW: Screen Resolution Pixels in Width, giving the number of pixels in width for the screen. This field takes a value 0 if no information/update is required to be sent.
- SRPH: Screen Resolution Pixels in Height, giving the number of pixels in height for the screen. This field takes a value 0 if no information/update is required to be sent.
- BLS: Battery Level Status. The following values are defined:

  * 0 if no information/update is required to be sent
  * 1= fading,
  * 2= weak,
  * 3= medium,
  * 4 = high,
  * 5= very high.
- SSz: Screen Size of the device. The following values are defined:

  * 0 if no information/update is required to be sent,
  * 1= very small,
  * 2= small,
  * 3= medium,
  * 4= big,
  * 5= very big.
- EnvNs: Environment noise level. The following values are defined:

  * 0 if no information/update is required to be sent
  * 1= very small,
  * 2= small,
  * 3= medium,
  * 4= noisy,
  * 5= very noisy.
- EnvLt: Environment light level. The following values are defined:

  * 0 if no information/update is required to be sent
  * 1= poor,
  * 2= dim,
  * 3= good,
  * 4= bright,
  * 5= very bright
- MStat: User activity and mobility status. The following values are defined:

  * 0 if no information/update is required to be sent
  * 1 = static,
  * 2 = weak mobility (e.g., walking),
  * 3= regular mobility (e.g., running),
  * 4 = high mobility (e.g., in train)
- Location: Gives the user location. This field takes a value 0 if no information/update is required to be sent.

The FEEDBACK option response has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Option code=TBA|Reserved Option|  Option Length= 16            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     PCP Server Notification                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               SRPW                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               SRPH                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   BLS |   SSz |EnvNs|EnvLt|MStat|      Location               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: FEEDBACK Option Response

The field PCP Server Notification shows the response status for the sent request:

- o 0= request cannot be satisfied,
- o 1= the request can be partially satisfied,
- o 2= the request can be satisfied,
- o 3= the request can be fully satisfied.

The content of the remaining fields are echoed from the request.

6. Security Consideration

The security consideration for PCP in RFC 6887 [RFC6887] and the PCP client authentication [PCPAuthentication] are sufficient to ensure security and host authorization for the proposed PCP extension in this document.

7. IANA Consideration

IANA is requested to assign a new PCP option called FEEDBACK in the IANA registry for PCP [pcp-iana].

8. Acknowledgements

Thanks for colleagues from Intel who provided valuable comments on this draft: Jeffrey Foerster, Somayazulu Srinivasa, Wu-Chi Feng and Muthaiah Venkatachalam.

Thanks also for Saso Stojanovski from Intel for the disucssion and exchange on PCP within 3GPP.

9. References

9.1. Normative References

[I-D.ietf-pcp-server-selection]
Boucadair, M., Penno, R., Wing, D., Patil, P., and T. Reddy, "PCP Server Selection", draft-ietf-pcp-server-selection-01 (work in progress), May 2013.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC6887] Wing, D., "Port Control Protocol (PCP)", RFC 6887, April 2013.

9.2. Informative References

[Flowdata]
Wing, D., Penno, R., and T. Reddy, "PCP Flowdata Option", draft-wing-pcp-flowdata-00 (work in progress), July 2013.

[I-D.ietf-httpbis-http2]
Belshe, M., Peon, R., Thomson, M., and A. Melnikov, "Hypertext Transfer Protocol version 2.0", draft-ietf-httpbis-http2-07 (work in progress), October 2013.

[PCPAuthentication]
Wasserman, M., Hartman, S., and D. Zhang, "Port Control Protocol (PCP) Authentication Mechanism", draft-ietf-pcp-authentication-02 (work in progress), October 2013.

[PCPProxy]
Boucadair, M., Penno, R., and D. Ding, "Port Control Protocol (PCP) Proxy Function", draft-ietf-pcp-proxy-03 (work in progress), June 2013.

[RFC2616] Fielding, R., "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC3840] Rosenberg, J., "Indicating User Agent Capabilities in Session Initiation Protocol (SIP)", RFC 3840, August 2004.

[RFC4566] Handley, M., "SDP: Session Description Protocol", RFC 4566, July 2006.

[ServiceFunctions]
Liu, W., Li, H., Huang, O., Boucadair, M., Leymann, N., Cao, Z., and J. Hu, "Service Function Chaining Use Cases",

draft-liu-sfc-use-cases-00 (work in progress), October 2013.

Appendix A. Existing Protocols of Relevance to User’s Feedback Information and their Limitations

Some existing protocols can convey devices characteristics and information on the user, however with limited applicability. Examples are:

- o SIP [RFC3840]

  * Provides a specification for capabilities and characteristics in SIP User Agent (UA). Capabilities and characteristics information is carried as parameters of the Contact header field and can be used within REGISTER requests and responses, OPTIONS responses, and requests and responses creating dialogs (such as INVITE).

  * SIP limitation for the explained use-cases: i) the need of adding the SIP protocol stack in video streaming servers, ii) SIP does not rely on network entities and is mainly an application specific protocol, and iii) SIP is not appropriate for "live" real-time control with no network priority to SIP controls.

  * Other Signalling protocols such as IAX are used to establish media sessions.

- o HTTP 1.1 [RFC2616] and HTTP2.0 [I-D.ietf-httpbis-http2]

  * HTTP can incorporate information on the device and the user in its headers (e.g. Accept or Use-Agent headers), however this will not be a generic solution to any underlying transport protocol.

  * Also, HTTP by its nature is a stateless protocol and activating HTTP proxy functionality impacts the performance of the network which limits the communication through proxies.

- o SDP [RFC4566]

  * SDP is meant to provide a standard representation for session description without incorporating a transport protocol, without a focus on user’s feedback information

* SDP is used for the session negotiation at the beginning of the session and so for the devices characteristics and the user information could not be directly signaled in real-time

* SDP uses protocols as SIP and RTSP that are not intercepted by middle nodes in the network which limits its applicability.

* SDP is not used in some non-SIP deployement contexts.

Authors' Addresses

Hassnaa Moustafa
Intel Corporation
Hillsboro, OR 97124
USA

EMail: hassnaa.moustafa@intel.com

Danny Moses
Intel Corporation

EMail: danny.moses@intel.com

Mohamed Boucadair
France Telecom
Rennes 35000
France

EMail: mohamed.boucadair@orange.com