

Descubrimiento de PCE inter-AS: una aportación a la computación de LSP en sistemas multidominio

M. Domínguez-Dorado¹, José-Luis González-Sánchez¹, J. Domingo-Pascual²

¹ Universidad de Extremadura

² Universidad Politécnica de Cataluña

mdomdor@unex.es, jlgs@unex.es, jordi.domingo@ac.upc.es

Abstract — El proceso de cálculo de las rutas que debe seguir el tráfico a través de Internet ha aumentado notablemente su complejidad en los últimos años. Hoy en día, este proceso está sujeto a la aplicación de múltiples restricciones relativas, entre otros aspectos, a la ingeniería de tráfico, a la gestión de recursos, a la calidad de los servicios ofrecidos al usuario final y a la seguridad y robustez de las comunicaciones. La aplicación de todas estas restricciones ha provocado que los elementos de red encargados de dicha tarea sean cada vez más complejos, de tal forma que, en algunos casos, el tiempo y los recursos dedicados al cálculo de rutas pueden afectar a la tarea principal del nodo: el reenvío de tráfico.

La arquitectura PCE (*Path Computation Element*) está siendo desarrollada para liberar de esta carga a los nodos en el contexto de la computación de caminos basada en restricciones para MPLS (*Multiprotocol Label Switching*). Aunque las investigaciones sobre PCE avanzan a buen ritmo, lo cierto es que aún quedan determinados aspectos por resolver. Como contribución a su desarrollo, en este trabajo se proporciona un mecanismo llamado, PILEP (*Procedure for Interdomain Location of External PCEs*), que permite el descubrimiento dinámico de elementos de computación de rutas en sistemas interdominio, haciendo uso de los protocolos de encaminamiento existentes.

Index terms — BGP, Descubrimiento inter-AS de PCE, MPLS, OSPF, QoS, TE.

I. INTRODUCCIÓN

Desde la primera implementación del protocolo RIP (*Routing Information Protocol*), los mecanismos de encaminamiento han avanzado notablemente. Actualmente, Internet integra multitud de redes heterogéneas donde coexisten múltiples tecnologías: IP (*Internet Protocol*), Ethernet, ATM (*Asynchronous Transfer Mode*), MPLS, etc. Cada una de estas tecnologías tiene sus propias características que afectan, en mayor o menor medida, a la capacidad de la red para proporcionar servicios de calidad a los usuarios. La ingeniería de tráfico (*TE – Traffic Engineering*) [1] y la QoS (*Quality of Routing*) [2] son dos disciplinas existentes encaminadas a aumentar esta capacidad; permiten modificar la naturaleza de Internet orientando a conexión aquellas redes que no tienen esta característica, realizar reservas de recursos, computar caminos de respaldo, proporcionar tolerancia a fallos, calcular rutas encaminadas a proporcionar calidad de servicio a las comunicaciones, etc.

MPLS (*Multiprotocol Label Switching*) [3] permite que este trabajo sea más sencillo al integrar las diferentes tecnologías y

la gestión de sus respectivos planos de control. La confluencia de todas estas nuevas circunstancias, ha repercutido en el proceso de cálculo de rutas; mientras hace algunas décadas el objetivo principal era que el tráfico llegara al destino (cuándo y cómo fuera), hoy en día en el proceso de cálculo de caminos se deben tener en cuenta nuevos aspectos que nunca antes habían sido necesarios. Por ejemplo, la aplicación de políticas de encaminamiento [4], la minimización de costes económicos, la maximización de los recursos de la red, la gestión del tráfico circulante, la adaptabilidad y recuperación de la red ante fallos o el cálculo de rutas disjuntas [5], entre otros.

Debido a la aplicación de todas las restricciones comentadas, el proceso de cómputo de caminos es ahora mucho más complejo y consume más recursos. Tradicionalmente, los nodos encargados de la clasificación y del reenvío de tráfico (usualmente los nodos de entrada a la red MPLS) han sido los que han realizado el cómputo de rutas, pero, en la actualidad, este trabajo es demasiado costoso. Hace varios años, el IETF (*Internet Engineering Task Force*) comenzó una serie de investigaciones para desarrollar una nueva tecnología encaminada a liberar a los nodos MPLS de las tareas de computar LSP (*Label Switched Path*) de tal forma que su labor central fuese el establecimiento del LSP y el reenvío de tráfico; el resultado de estos trabajos ha sido la definición de la arquitectura PCE (*Path Computation Element*) [6].

El objetivo general de nuestro trabajo es contribuir al desarrollo de la arquitectura PCE en aquellos entornos, como el interdominio, donde aún existen aspectos por resolver. Para ello aportamos PILEP (*Procedure for Interdomain Location of External PCEs*), un mecanismo dinámico de localización de elementos PCE en entornos inter-AS (*inter Autonomous Systems*).

El resto de este documento está organizado de la siguiente forma: en el segundo apartado se realiza una breve descripción de la arquitectura PCE; el tercer apartado resalta los motivos por los cuales es deseable diseñar y desarrollar un mecanismo dinámico de descubrimiento de elementos PCE en entornos interdominio; en la cuarta sección presentamos PILEP, nuestra propuesta para el descubrimiento de elementos PCE inter-AS; la sección cinco muestra un ejemplo de su aplicación; finalmente, en el apartado seis comentamos las conclusiones y el trabajo futuro.

II. DESCRIPCIÓN DE LA ARQUITECTURA PCE

La arquitectura PCE está siendo desarrollada actualmente, por lo que la mayor parte de los RFC publicados por el IETF contienen definiciones de requisitos generales y descripciones de la arquitectura.

En la arquitectura PCE más básica deben existir al menos tres elementos principales (Fig. 1). PCE es el elemento encargado de calcular rutas; PCC (*Path Computation Client*) es el cliente que solicitará al PCE el cómputo de caminos; y PCEP (*Path Computation Element communication Protocol*) [7], [8], es el protocolo de comunicaciones mediante el cual se comunican PCE y PCC. Generalmente los nodos LER (*Label Edge Router*), como nodos de entrada a la red MPLS, serán los que actuarán como PCC.

Aunque a primera vista parece un modelo simple, existen diversas dificultades colaterales que surgen al integrar esta nueva arquitectura en un sistema autónomo con otras tecnologías heredadas, problemas que son objeto de estudio por parte de investigadores de todo el mundo; por ejemplo, la relación de PCE con los protocolos de encaminamiento IGP (*Interior Gateway Protocol*) y EGP (*Exterior Gateway Protocol*) o el suministro de información de ingeniería de tráfico a los PCE por parte de los protocolos existentes.

En la arquitectura PCE, un AS (*Autonomous System*), o dominio MPLS, debe contar con uno o más elementos PCE destinados a computar los LSP dentro de él. Cada nodo que desee iniciar el establecimiento de un LSP debe actuar como PCC; por tanto, al menos los nodos LER deben ser nodos PCC, puesto que sobre ellos recae la responsabilidad de establecer LSP a lo largo del sistema autónomo. Además de ellos, otros nodos intermedios podrían necesitar actuar como PCC si actualmente participan en la restauración de LSP y necesitan calcular rutas.

Cuando un flujo llega al LER de entrada al dominio, este LER actuará como PCC y solicitará al PCE que compute un LSP desde él hasta el LER de salida del dominio. Para ello, utilizará el protocolo PCEP, que proporciona suficiente funcionalidad para permitir al LER/PCC de entrada realizar la solicitud de cómputo. Ésta, irá acompañada de un conjunto de restricciones que el PCE deberá tener en cuenta durante el cálculo. El PCE calculará la ruta basándose en la información contenida en la TED (*Traffic Engineering Database*), que es una base de datos que contendrá el grafo de estado de la red y cualquier información adicional que pueda ser de utilidad.

Cada PCE está asociado a una TED que es actualizada

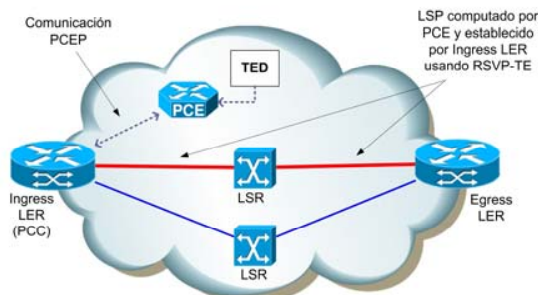


Fig. 1. Componentes básicos en la arquitectura PCE.

periódicamente por los IGP (*Interior Gateway Protocol*) o por otros medios alternativos que se pudieran definir. Una vez que la ruta ha sido calculada, el PCE utilizará de nuevo PCEP para enviar la respuesta al LER/PCC que realizó la solicitud.

En la definición de requisitos para el protocolo PCEP [7] se especifica que la ruta calculada, incorporada en la respuesta al LER/PCC, debe ser directamente transformable en un objeto ERO (*Explicit Routing Object*) de RSVP-TE (*Resource Reservation Protocol – Traffic Engineering*) [9], de tal forma que el LER/PCC sea capaz de comenzar el establecimiento del LSP utilizando RSVP-TE y dicho ERO.

Este es el modo de funcionamiento más básico de la arquitectura PCE (computación simple). PCE permite la existencia de otras situaciones que pueden resultar más complejas de coordinar, como se puede ver en la Fig. 2. Por ejemplo, puede haber más de un elemento PCE encargado de calcular LSP completos dentro del AS. En ese caso, un LER/PCC tendrá la posibilidad de seleccionar de entre varios PCE, aquel que más se ajusta a sus necesidades. Esta situación obliga a que el LER/PCC conozca cierta información sobre las capacidades de los PCE disponibles, a fin de poder realizar el proceso de selección basándose en un criterio razonable.

Otro ejemplo es aquel en el que existe más de un PCE en el AS, pero cada uno de ellos encargado de calcular segmentos de ruta sobre un área concreta de la red (computación múltiple). En este caso, los PCE se verán obligados a colaborar entre sí para calcular los respectivos segmentos, ensamblarlos y devolver al LER/PCC la ruta completa que éste solicitó.

Por tanto, existen ciertas circunstancias en las que un PCE debe actuar como PCC de cara a otro PCE. La situación más compleja posible se da cuando las dos situaciones anteriores coinciden y la ruta solicitada por el LER/PCC sobrepasa los límites del AS local. En esta situación es donde la arquitectura PCE debe contar con todos los mecanismos posibles para superar las barreras que, tradicionalmente, han existido en el encaminamiento interdominio [10], [11] en relación con la ingeniería de tráfico y con MPLS como, por ejemplo, la visión parcial de la topología, la disponibilidad limitada de información de ingeniería de tráfico, el encaminamiento basado en políticas, la unicidad en el cálculo de rutas, la intimidad de los AS, la seguridad o los mecanismos de recuperación de la red ante fallos.

III. ARGUMENTOS PARA EL DISEÑO DE UN MECANISMO DE DESCUBRIMIENTO DE ELEMENTOS PCE EN INTER-AS

Independientemente del caso que estemos tratando, el primer paso que debe tener lugar es el descubrimiento de los PCE existentes por parte de los potenciales PCC. En el caso más simple, un LER/PCC necesitará saber a qué PCE enviar su solicitud de cálculo de ruta; en el caso de PCE que deban colaborar entre sí cada PCE deberá conocer la existencia del resto de PCE con los que deberá colaborar.

Para calcular LSP en entornos inter-AS, es necesaria la colaboración entre PCE, por lo que el proceso de descubrimiento cobra, en esta situación, una especial relevancia.

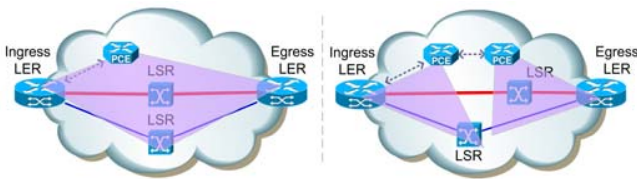


Fig. 2. Computación simple (izda.) y computación múltiple (dcha.).

En los siguientes subapartados se detallan los mecanismos existentes para el descubrimiento de elementos PCE en el interior de un dominio y la dificultad para extrapolar estos mecanismos a un entorno interdominio.

A. Descubrimiento intra-AS basado en IGP

En [12], se especifica que un PCC debe conocer la existencia de un PCE utilizando cualquier método excepto el *broadcast*. Además, impone otros requisitos como el hecho de que un PCE pueda ser descubierto con más o menos nivel de detalle por distintos PCC o que se preserve la intimidad del AS en el caso de un descubrimiento interdominio. Todo ello está orientado a disponer de un método flexible que permita a los AS establecer distintas políticas de descubrimiento.

Conforme a estos requisitos, existen dos propuestas que utilizan los mecanismos de inundación de los protocolos OSPF (*Open Shortest Path First*) [13] e IS-IS (*Intermediate System to Intermediate System*) [14] para difundir la existencia de un PCE en el interior de un sistema autónomo, pudiendo ser detectado por los PCC interesados. En [13] se define una nueva terna TLV (*Type Length Value*) para OSPF llamada PCED (*PCE Discovery*) que puede ser utilizada dentro de un mensaje de tipo RI-LSA (*Routing Information – Link State Advertisement*) de OSPF [15] en conjunción con sus extensiones para ingeniería de tráfico [16]. Los mensajes de tipo RI-LSA se intercambian periódicamente entre nodos OSPF para mantener actualizadas las bases de datos sobre la topología de la red. Estas bases de datos, propias de cada nodo OSPF, son utilizadas para calcular el siguiente salto hacia el destino para el tráfico saliente.

Este método de descubrimiento, añade a estos mensajes información adicional (el TLV PCED) de tal forma que, a la vez que se difunde información sobre el estado de los enlaces, se difunda también información sobre la existencia de un PCE. Para ello, el elemento PCE que desee ser descubierto debe participar también en el funcionamiento del protocolo OSPF. El TLV PCED (Fig. 3) está constituido por un conjunto no ordenado de sub-TLV. Estos sub-TLV contienen información suficiente para que los elementos PCC y PCE dentro del ámbito de inundación de OSPF sean capaces de detectar los PCE existentes y elegir el adecuado en cada momento. Siempre informará sobre la dirección IP que debe usar un PCC para alcanzar al PCE anunciado (sub-TLV PCE-ADDRESS) y el ámbito de trabajo de dicho PCE (sub-TLV PCE-SCOPE), por ejemplo, inter-AS, inter-área, inter-capa, etc.

Además, TLV PCED puede incorporar de forma opcional información más detallada sobre el PCE anunciado: PCE-DOMAIN y NEIG-PCE-DOMAIN permiten indicar al PCC el conjunto de áreas o sistemas autónomos sobre los cuales el

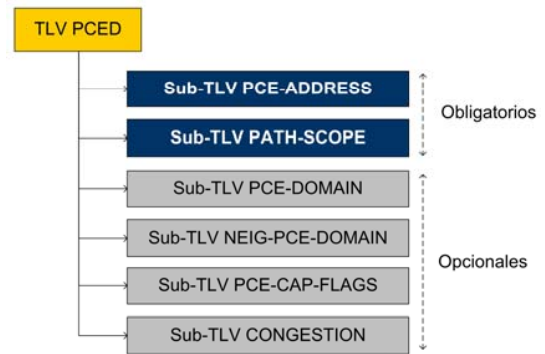


Fig. 3. Formato del TLV PCED de OSPF-TE.

PCE tiene capacidad de calcular LSP; PCE-CAP-FLAGS incorpora información sobre las capacidades del PCE que son potencialmente útiles para seleccionar entre diversos PCE candidatos (como por ejemplo la bidireccionalidad, o la priorización de peticiones).

Finalmente, CONGESTION se utiliza para indicar si el PCE que se está anunciando está experimentando congestión o no. Empleando este método, un PCE que participe de OSPF puede anunciarse a sí mismo a aquellos PCE o PCC susceptibles de hacer uso de él, mediante los mecanismos de inundación IGP tradicionales. Cada PCC o PCE en el ámbito de inundación OSPF descubrirá, evitando el *broadcast*, la existencia del PCE y suficiente información para realizar un proceso de selección de PCE adecuado (Fig. 4). Este método resulta especialmente ventajoso, puesto que aprovecha mecanismos preexistentes en el AS para difundir la existencia y el estado de un PCE en el interior de un AS, con modificaciones mínimas de los protocolos existentes.

B. Necesidad de mecanismos de descubrimiento inter-AS

Pese a sus bondades, el mecanismo anterior no puede ser extrapolado a un entorno interdominio. Los protocolos IGP en los que se apoya, OSPF e IS-IS, tienen un ámbito de actuación limitado que se circunscribe al interior de un sistema autónomo y, por tanto, no son válidos para transportar información sobre elementos PCE entre AS adyacentes. Aunque en [12] se contemplan los requisitos para ello, aún no se ha desarrollado un mecanismo dinámico y automático para que un PCE perteneciente a un sistema autónomo pueda conocer la existencia y las capacidades de otro elemento PCE en un sistema autónomo adyacente.

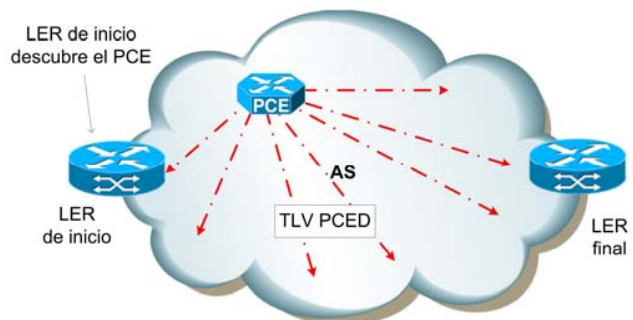


Fig. 4. Uso del TLV PCED para el descubrimiento de PCE vía OSPF.

En la Fig. 5, se observa un ejemplo de sistema interdominio donde es necesaria la colaboración entre PCE para computar un LSP desde el LER de inicio (en AS1) hasta el LER final (en AS2). Cada elemento PCE tiene la capacidad de computar rutas sólo en el interior de su respectivo AS.

En este entorno, el elemento PCE de AS1 debe poder descubrir de forma dinámica que existe un PCE en AS2 con el que puede colaborar para ofrecer al LER/PCC de inicio la ruta solicitada hacia el LER final. Por el momento, la única solución existente consiste en que se suscriban acuerdos de colaboración entre los sistemas autónomos implicados y se configuren manualmente las relaciones entre los distintos elementos PCE; sin embargo, es deseable que este proceso se pueda realizar de forma dinámica y automática.

En este trabajo, aportamos un mecanismo que cubre este espacio, aún por solucionar, permitiendo a un PCE descubrir la existencia de otro elemento PCE en un AS adyacente. Para ello, extenderemos los mecanismos de descubrimiento de elementos PCE propuestos por el IETF para el descubrimiento en el interior de un AS.

IV. PROPUESTA DE MECANISMO DE DESCUBRIMIENTO DE ELEMENTOS PCE EN ENTORNOS INTER-AS

El descubrimiento de un PCE en el interior de un AS es, en su concepción, distinto al descubrimiento de un PCE en entornos interdominio. En el caso intra-AS, se intenta hacer visible un PCE para los posibles PCC del AS, sin importar demasiado las condiciones en las que este descubrimiento se produce. En el caso del interdominio, se debe hacer visible un PCE perteneciente a un AS, a otro PCE perteneciente a un AS adyacente (que actuará de PCC) para colaborar en el cómputo de un LSP interdominio. Entra en juego, por tanto, el problema de preservar la intimidad de los AS.

Nuestra propuesta, PILEP, extiende [13] añadiendo nuevas características para permitir a un conjunto de PCE en sistemas autónomos externos conocer la existencia de elementos PCE en el AS local. Para este propósito, ha sido necesario añadir unas extensiones mínimas, tanto a OSPF como a BGP (*Border Gateway Protocol*), pues son parte actora dentro del mecanismo de descubrimiento inter-AS.

A. Extensión de OSPF-TE: sub-TLV SET-VISIBLE-TO

La primera extensión de protocolo necesaria para el funcionamiento de nuestra propuesta es del protocolo OSPF-TE. Éste, debe conservar la capacidad de difundir elementos PCE en el interior de un AS pero, además, debe poder notificar a los ASBR (*Autonomous System Border Router*) la necesidad de que colaboren transportando el anuncio más allá del límite del sistema autónomo.

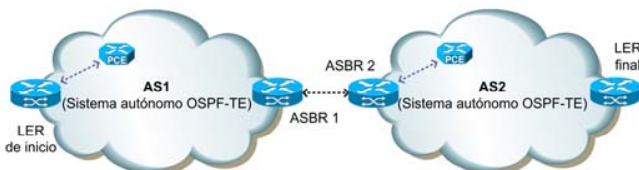


Fig. 5. Entorno con necesidad de descubrimiento de PCE inter-AS.

Para conseguir este objetivo, hemos diseñado una extensión (Fig. 6) del TLV PCED utilizado en los mecanismos de descubrimiento interior, añadiéndole una nueva terna sub-TLV llamada SET-VISIBLE-TO. Este nuevo sub-TLV es similar al TLV PCED, pero dirigido a sistemas autónomos adyacentes y no al sistema autónomo local.

SET-VISIBLE-TO es opcional y de tamaño variable dentro del TLV PCED. Puede existir más de una instancia de él dentro del TLV PCED lo que permite disponer de diferentes conjuntos de información dirigida a los distintos AS adyacentes. Este modo de funcionamiento permite que los PCE de los sistemas autónomos adyacentes puedan conocer la existencia del PCE que se anuncia, en diferentes grados, habilitando a los AS para definir distintas políticas de descubrimiento dependiendo del AS objetivo del anuncio. SET-VISIBLE-TO tiene un formato (Fig. 7) similar a PCED. Si un sub-TLV SET-VISIBLE-TO se encuentra presente en un TLV PCED, deberá incluir, al menos, cuatro sub-sub-TLV que son obligatorios: SVT-PCE-ADDRESS cuyo significado es el mismo que PCE-ADDRESS en PCED; SVT-PATH-SCOPE, similar a PATH-SCOPE en PCED, SVT-PCE-DOMAIN, que indica el sistema autónomo al que pertenece el PCE que se está anunciando (y sobre el que podrá calcular LSP); y SVT-TARGET-AS (Fig. 8), que indica el identificador del AS a los que se debe hacer llegar la existencia del PCE que se está anunciando. Además de los sub-TLV obligatorios, SET-VISIBLE-TO puede transportar otros opcionales con el objeto de facilitar el proceso de selección de PCE: SVT-NEIG-DOMAIN, que contiene los identificadores de AS adyacentes sobre los cuales el PCE anunciado puede calcular LSP (al margen del AS local); y finalmente, SVT-PCE-CAP-FLAGS y SVT-CONGESTION con los mismos significados que PCE-CAP-FLAGS y CONGESTION en PCED. Un resumen sobre el formato de SET-VISIBLE-TO se puede ver en la Tabla 1.

Utilizando SET-VISIBLE-TO, un encaminador BGP que participe en OSPF es capaz de entender que un TLV PCED debe traspasar las fronteras del AS local y ser retransmitido a uno o más AS adyacentes. Además, conocerá también qué información debe mostrar a cada uno de estos sistemas autónomos, como se requiere en [12] para el descubrimiento en entornos interdominio.

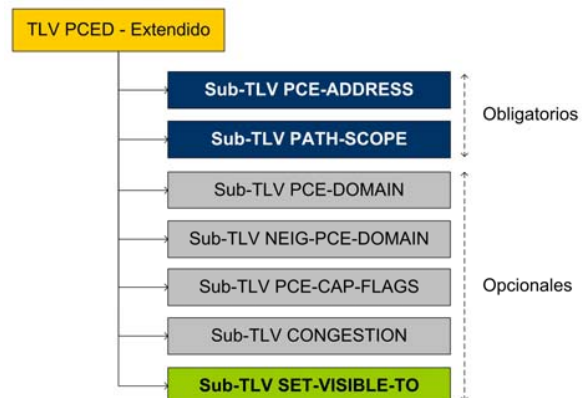


Fig. 6. Ubicación del sub-TLV SET-VISIBLE-TO en PCED.

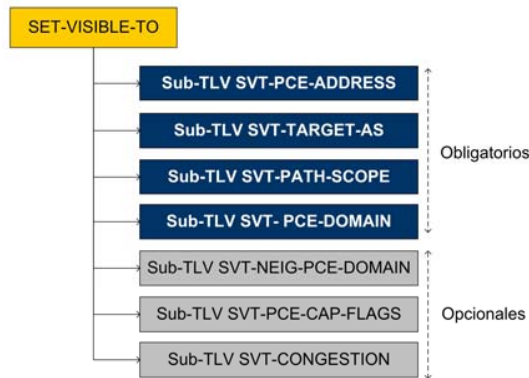
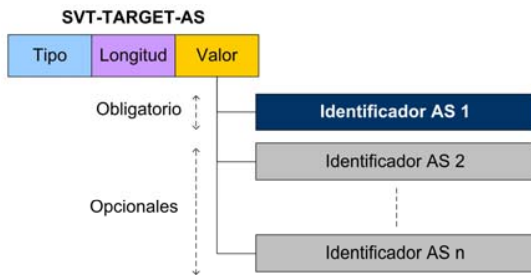


Fig. 7. Estructura del sub-TLV SET-VISIBLE-TO.



Tipo: 1 octeto. Tipo de sub-sub-TLV. A definir por IANA.
Longitud: 1 octeto. Nº de entradas existentes en el campo Valor.
Valor: secuencia de Identificadores de AS. Su tamaño en octetos es de 2 x Longitud.

Fig. 8. Formato de SVT-TARGET-AS.

TABLA 1
 RESUMEN DEL FORMATO DEL SUB-TLV SET-VISIBLE-TO

Sub-sub-TLV	Oblig.	Explicación
SVT-PCE-ADDRESS	✓	Dirección IP del elemento PCE anunciado.
SVT-TARGET-AS	✓	Identificador/es de el/los sistema/s autónomo/s que debe/n conocer la existencia del PCE que se anuncia.
SVT-PATH-SCOPE	✓	Inter-AS, inter-capa, inter-área, como se especifica en [13] para PATH-SCOPE.
SVT-PCE-DOMAIN	✓	Identificador del AS al que pertenece el PCE anunciado. Igual que PCE-DOMAIN definido en [13], pero sin la posibilidad de expresar identificador de área; sólo de AS.
SVT-NEIG-PCE-DOMAIN	×	Formato y significado idénticos NEIG-PCE-DOMAIN definido en [13], pero sin la posibilidad de expresar identificador de área; sólo de AS.
SVT-PCE-CAP-FLAGS	×	Idéntico formato y significado que PCE-CAP-FLAGS, como se detalla en [13].
SVT-CONGESTION	×	Idéntico formato y significado que CONGESTION, como se detalla en [13].

B. Extensión de BGP: atributo AS_PCE

Una vez implementados los mecanismos para que un ASBR entienda que debe hacer llegar el anuncio de elementos PCE a otros sistemas autónomos adyacentes, el siguiente paso consiste en dotar a BGP de los elementos necesarios para que

esta información pueda traspasar, efectivamente, los límites del AS local. Esta es la segunda extensión necesaria para el mecanismo de descubrimiento de PCE inter-AS.

BGP define la utilización de atributos de ruta [17] para caracterizar el camino que debe seguir el tráfico en un sistema interdominio en su transcurso hacia el destino. Además, especifica el mecanismo para la utilización de atributos opcionales que no tienen por qué ser soportados por todas las implementaciones de BGP. En PILEP se aprovecha esta característica de BGP para hacer llegar la información de descubrimiento de elementos PCE de un sistema autónomo a otro adyacente.

Proponemos un nuevo atributo de ruta BGP, opcional, no transitivo, llamado AS_PCE (Fig. 9). Al ser un atributo opcional y no transitivo, no existe la obligación de que los ASBR que no implementen esta propuesta deban entenderlo; y además, el atributo no llegará más allá del AS adyacente al que se desea que llegue la información. El contenido de este atributo de ruta coincide exactamente con el del sub-TLV SET-VISIBLE-TO, a excepción de la lista de AS destino de la información que, en el atributo AS_PCE, se ha eliminado.

Un nodo BGP/OSPF-TE que implemente PILEP, trabajará de la siguiente forma: si detecta vía OSPF-TE un TLV PCED que incorpora al menos un sub-TLV SET-VISIBLE-TO, analizará si mantiene acuerdos de *peering* con alguno de los AS especificados en SVT-TARGET-AS.

Para cada uno de los AS especificados en SVT-TARGET-AS con los que mantenga relaciones de *peering*, creará un atributo de ruta AS_PCE que será adjuntado a las rutas hacia el AS local que se encuentren en las tablas de encaminamiento BGP (Adj-RIB-Out) [17]. Si no existiese ninguna ruta hacia el AS local, se creará una y se le adjuntará el atributo. De esta forma, se consigue que el anuncio de un PCE sobrepase los límites del AS utilizando los mensajes UPDATE del protocolo de encaminamiento BGP.

C. Gestión de mensajes UPDATE en el ASBR receptor

Debido al proceso de selección de rutas llevado a cabo por los nodos BGP, existe la posibilidad de que la ruta que incorpora el atributo AS_PCE sea desestimada en dicho proceso. Por si esto ocurre, es necesario poder rescatar la información de ese atributo como paso previo al proceso de selección.

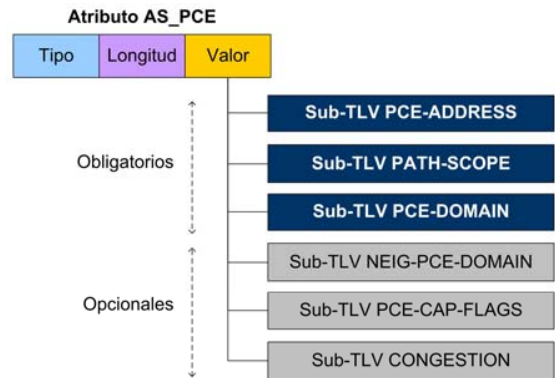


Fig. 9. Estructura del atributo de ruta AS_PCE.

El mecanismo de gestión de mensajes UPDATE entrantes se muestra en la Fig. 10. Cuando un nodo BGP/OSPF-TE que implemente PILEP reciba una ruta de un sistema autónomo adyacente, deberá comprobar si contiene un anuncio de PCE (si contiene un atributo AS_PCE).

En caso afirmativo, desviará una copia de la misma para ser procesada por el módulo OSPF-TE extendido, ya que OSPF-TE es el protocolo encargado de difundir información de encaminamiento en el interior del sistema autónomo objetivo.

La ruta, además, será procesada paralelamente por el módulo BGP de la forma habitual.

Por su parte, y al margen del proceso de selección de rutas llevado a cabo por BGP, el módulo extendido OSPF-TE tomará el atributo AS_PCE y lo transformará en un TLV PCED para ser difundido junto con los mensajes RI-LSA en el interior del sistema autónomo objetivo. Esta transformación es directa, puesto que el campo "Valor" del atributo AS_PCE guarda el mismo formato que el campo "Valor" del TLV PCED. Esta transformación y su difusión, cierran el ciclo del mecanismo de descubrimiento. El TLV PCED original habrá sufrido diversas transformaciones en su trayecto: de ser un PCED extendido con un sub-TLV SET-VISIBLE-TO a un atributo AS_PCE de un mensaje UPDATE de BGP para, finalmente, convertirse en un TLV PCED tradicional en el sistema autónomo destino.

V. EJEMPLO DE FUNCIONAMIENTO

El proceso propuesto involucra mensajes de diversas tecnologías. Para ayudar a entender el funcionamiento global de PILEP en un caso concreto, se aplicarán a continuación todas las fases del mecanismo, sobre el ejemplo que se especificó en la Fig. 5: un sistema interdominio simple con dos PCE y necesidad de colaboración entre ellos.

A. Difusión del elemento PCE de AS2 intra-AS

El primer paso que debe seguir el elemento PCE de AS2 es difundir su existencia utilizando para ello el TLV PCED extendido y el mecanismo basado en OSPF-TE definido por el IETF. PCED incorporará en este caso un sub-TLV SET-VISIBLE-TO donde se especificará su dirección IP, el número de sistema autónomo que identifica a AS1, indicará que puede colaborar en el cómputo de LSP inter-AS y que puede hacerlo sobre AS2.

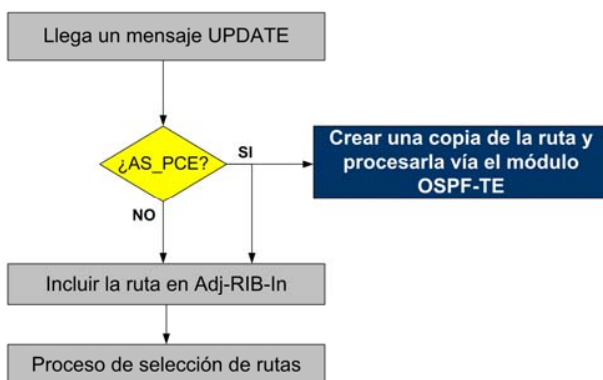


Fig. 10. Gestión de mensajes UPDATE con atributo AS_PCE.

Aunque puede incorporar información adicional, para este ejemplo supondremos que no será así. Este PCED extendido se adjuntará al siguiente mensaje RI-LSA y se difundirá en el interior de AS2 (Fig. 11).

B. Detección del PCE y transmisión inter-AS

ASBR2 es un nodo que participa tanto en BGP como en OSPF-TE en el AS2. Además, implementa la propuesta expresada en este documento. Por tanto, con la llegada del mensaje RI-LSA (y el PCED extendido que incorpora) detecta que hay un elemento PCE en su propio AS; simultáneamente, debido a la existencia del sub-TLV SET-VISIBLE-TO, entiende que el anuncio debe traspasar la frontera del AS y, como él es un ASBR, procesa SET-VISIBLE-TO. Al hacerlo, descubre que el anuncio debe llegar a AS1, con el que mantiene acuerdos de *peering*. Esto provoca que cree un atributo AS_PCE a partir de la información contenida en SET-VISIBLE-TO y lo incorpore a las rutas hacia el AS local que debe transmitir a AS1. Con el siguiente envío de un mensaje UPDATE, el atributo AS_PCE llegará a AS1 desde AS2, vía BGP (Fig. 12).

C. Gestión del mensaje UPDATE en ASBR 1

ASBR1, al igual que ASBR2, participa tanto en OSPF como en BGP e implementa PILEP. Recibe un mensaje UPDATE de ASBR y analiza sus atributos. Descubre el atributo AS_PCE y, por tanto, lo envía al módulo OSPF-TE extendido para su tratamiento, a la vez que la ruta queda en Adj-RIB-in (tabla de encaminamiento BGP en crudo) para el proceso de selección de rutas.

D. Difusión del elemento PCE intra-AS en AS1

El módulo OSPF-TE de ASBR1 entiende que debe transformar AS_PCE en un TLV PCED tradicional y adjuntarlo al siguiente mensaje RI-LSA para que la información sea difundida a lo largo de AS1 (Fig. 13). Con esta difusión, el elemento PCE de AS1 descubrirá la existencia del elemento PCE de AS2. Además, descubrirá que dicho elemento pertenece a AS2, que puede colaborar con él en el cómputo de un LSP interdominio, que puede realizar el cómputo sobre AS2 y sabrá su dirección IP, con lo cual podrá establecer una sesión PCEP entre ambos (Fig. 14).

El elemento PCE de AS1 decidirá con qué PCE de AS2 colaborar para el cómputo de la ruta que el LER de inicio le ha solicitado. En este ejemplo no hay más PCE candidatos, por lo que utilizará el único elemento PCE de AS2. Tras los pasos anteriores, finaliza el proceso de descubrimiento y comienza el modo de operación normal de PCE.

Hasta este punto, PILEP cubre los requisitos de seguridad y de confidencialidad expresados en [12], permitiendo especificar el conjunto de AS a los que notificar la existencia del PCE y la definición de un conjunto de datos relativos a dicho PCE distinto para cada AS destino del anuncio. Adicionalmente, gracias a PILEP, un PCE puede ser anunciado con distintas direcciones IP a cada AS objetivo y esta característica puede ser utilizada por los administradores de la red para establecer un control de acceso basado en la dirección IP usada para conectar al PCE anunciado.

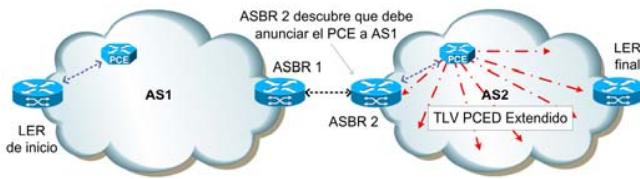


Fig. 11. Utilización del TLV PCED Extendido en inter-AS.



Fig. 12. Utilización del atributo de ruta AS_PCE.

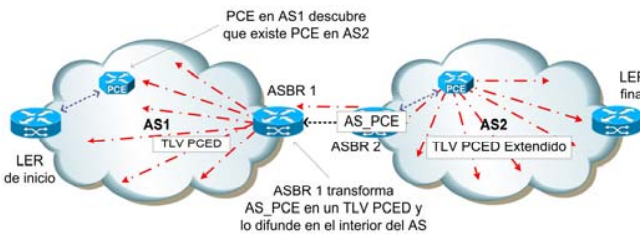


Fig. 13. Efecto de transformar AS_PCE en un TLV PCED.



Fig. 14. Computación distribuida de LSP inter-AS mediante PCE.

El IETF propone otros mecanismos de seguridad y privacidad complementarios [8], aunque dichos mecanismos actúan en conjunción con el protocolo PCEP y no sobre el mecanismo de descubrimiento.

VI. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo de investigación, aún en curso, proponemos PILEP, un mecanismo para el descubrimiento dinámico de elementos PCE en un entorno interdominio. Este mecanismo está justificado por la necesidad de cooperación entre elementos PCE a la hora de computar un LSP que sobrepase los límites de un sistema autónomo y por la inexistencia de un mecanismo de descubrimiento alternativo.

Nuestra propuesta sigue la línea emprendida por el IETF al utilizar los mecanismos existentes en los protocolos de encaminamiento tradicionales para facilitar el proceso de descubrimiento. Se ha diseñado una nueva sub-TLV opcional que puede adjuntarse a la TLV PCE de OSPF-TE y un nuevo atributo opcional, no transitivo, AS_PCE, para incorporarlo a

las rutas anunciadas por BGP en un mensaje UPDATE. Las ventajas de PILEP se podrían resumir en los siguientes puntos:

- Cubre la necesidad de un mecanismo para descubrir elementos PCE en entornos interdominio.
- Permite el establecimiento de distintas políticas de descubrimiento según el AS objetivo.
- Realiza su función con cambios mínimos en los protocolos involucrados.

Dado que el trabajo aún está en curso, existen ciertas tareas que se han de acometer en un futuro como, por ejemplo, la evaluación de la bondad del mecanismo propuesto utilizando simulaciones (actualmente se están analizando distintos simuladores sobre los que implementar PILEP), o la extensión de la propuesta para ser usada en conjunción con IS-IS.

AGRADECIMIENTOS

Este trabajo está financiado, en parte, por la Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura, Proyecto AGILA-2, con código No. PR1A06145; y por el Ministerio de Industria, Turismo y Comercio, Proyecto MESEAS, con código FIT-350301-2007-14.

REFERENCIAS

- [1] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, X. Xiao. *Overview and Principles of Internet Traffic Engineering*. IETF RFC 3272. May, 2002.
- [2] E. Crawley, R. Nair, B. Rajagopalan, H. Sandick. *A Framework for QoS-based Routing in the Internet*. IETF RFC 2386. August, 1998.
- [3] E. Rosen, A. Viswanathan, R. Callon. *Multiprotocol Label Switching Architecture*. IETF RFC 3031. January, 2001.
- [4] C. K. Chau, R. Gibbens, T. G. Griffin. *Towards a Unified Theory of Policy-Based Routing*. INFOCOM 2006, pp. 1-12. April, 2006.
- [5] A. Sprintson, M. Yannuzzi, A. Orda, X. Masip-Bruin. *Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks*. INFOCOM 2007, pp. 1820-1828. May, 2007.
- [6] A. Farrel, J. P. Vasseur, J. Ash. *A Path Computation Element (PCE)-Based Architecture*. IETF RFC 4655. August, 2006.
- [7] J. Ash, J.L. Le Roux. *Path Computation Element (PCE) Communication Protocol Generic Requirements*. IETF RFC 4657. September, 2006.
- [8] J. P. Vasseur, J. L. Le Roux. *Path Computation Element (PCE) communication Protocol (PCEP)*. IETF Draft draft-ietf-pce-pcep-12.txt. Work in progress. March, 2008.
- [9] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow. *RSVP-TE: Extensions to RSVP for LSP Tunnels*. IETF RFC 3209. December 2001.
- [10] M. Yannuzzi, X. Masip-Bruin, O. Bonaventure. *Open issues in interdomain routing: a survey*. IEEE Network. Volume 9, Issue 6. November-December, 2005.
- [11] C. Pelsser, S. Uhlig, O. Bonaventure. *On the difficulty of establishing interdomain LSPs*. Proceedings of the IEEE Workshop on IP Operations and Management. October 2004.
- [12] J.L. Le Roux. *Requirements for Path Computation Element (PCE) Discovery*. IETF RFC 4674. October, 2006.
- [13] J. L. Le Roux, J.P. Vasseur, Y. Ikejiri, R. Zhang. *OSPF Protocol Extensions for Path Computation Element (PCE) Discovery*. IETF RFC 5088. January 2008.
- [14] J. L. Le Roux, J.P. Vasseur, Y. Ikejiri, R. Zhang. *IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery*. IETF RFC 5089. January 2008.
- [15] A. Lindem, N. Shen, J. P. Vasseur, R. Aggarwal, S. Shaffer. *Extensions to OSPF for Advertising Optional Router Capabilities*. IETF RFC 4970. July, 2007.
- [16] D. Katz, K. Kompella, D. Yeung. *Traffic Engineering (TE) Extensions to OSPF Version 2*. IETF RFC 3630. September, 2003.
- [17] Y. Rekhter, T. Li, S. Hares. *A Border Gateway Protocol 4 (BGP-4)*. IETF RFC 4271. January, 2006.