

RI-CUBE: Dotando al PCE de información abstracta de ingeniería de tráfico interdominio

M. Domínguez-Dorado

Departamento de Ingeniería de Sistemas Informáticos y
Telemáticos (DISIT)
Universidad de Extremadura
Cáceres, SPAIN
mdomdor@unex.es

J. Domingo-Pascual

Departamento de Arquitectura de Computadores (DAC)
Universitat Politècnica de Catalunya
Barcelona, SPAIN
jordi.domingo@ac.upc.edu

J. L. González-Sánchez

Departamento de Ingeniería de Sistemas Informáticos y
Telemáticos (DISIT)
Universidad de Extremadura
Cáceres, SPAIN
jlgs@unex.es

J. Carmona-Murillo

Departamento de Ingeniería de Sistemas Informáticos y
Telemáticos (DISIT)
Universidad de Extremadura
Cáceres, SPAIN
jcarmur@unex.es

Abstract— El cómputo de rutas en Internet se ha vuelto una tarea compleja y costosa. La arquitectura PCE (*Path Computation Element*) proporciona la funcionalidad necesaria para el cómputo de rutas interdominio en redes MPLS (*Multiprotocol Label Switching*) y GMPLS (*Generalized Multiprotocol Label Switching*). En este escenario, el cálculo de rutas interdominio se lleva a cabo mediante la cooperación entre PCEs. El PCE que requiera ayuda, utiliza un mecanismo de selección de PCEs colaboradores que podría tener en consideración el estado de la red y sus recursos. Este mecanismo es especialmente importante debido al impacto que tiene en el tiempo total necesario para computar una ruta interdominio completa. En este trabajo, aportamos un detallado estudio de la información de ingeniería de tráfico manejada por los IGP (*Interior Gateway Protocols*) más importantes y también un mecanismo para intercambiar esta información en entornos interdominio de forma que no se viole la privacidad de sobre la topología de la dominios afectados. Con esta información en su poder, un elemento PCE puede seleccionar un PCE exterior para colaborar, de forma efectiva y más precisa, minimizando el tiempo total necesario para calcular la ruta interdominio.

Keywords- Encaminamiento interdominio; ISIS-TE; MPLS-TE; OSPF-TE; selección de PCE.

I. INTRODUCCIÓN

Hoy en día, la tarea de cálculo de rutas en Internet es una tarea ardua; en este proceso se debe tener en cuenta un conjunto creciente de restricciones que cada vez son más complejas. Por ejemplo, se podría requerir que una ruta

proporcionara un cierto ancho de banda, que equilibrara la carga en la red o que minimizara el ancho de banda residual en todos los enlaces de la topología. Para ello, se hace uso de todas las extensiones de ingeniería de tráfico (*TE, Traffic Engineering*) [1]. Ésta, puede ser entendida como la capacidad de algunas tecnologías de monitorizar, medir, gestionar y modificar el comportamiento de las redes en funcionamiento para sacar el mayor provecho de ellas y para proporcionar a los flujos circulantes la calidad de servicio (*QoS, Quality of Service*) [2] esperada.

La TE en Internet se articula mediante un conjunto de extensiones de protocolos y tecnologías auxiliares existentes. Gracias a ellas, los ISP (*Internet Service Providers*) pueden satisfacer las restricciones necesarias cuando calculan rutas sobre sus propios dominios. Por tanto, es previsible que un ISP que implemente técnicas de TE pueda asegurar la calidad de sus servicios mejor que uno que no implemente este tipo de extensiones.

A. Importancia de la ingeniería de tráfico basada PCE

La arquitectura PCE (*Path Computation Element*) [3] permite obtener TE en redes MPLS (*Multiprotocol Label Switching*) [4] y GMPLS (*Generalized Multiprotocol Label Switching*) [5], tanto en intradominio como en interdominio [7].

Uno de los objetivos principales de esta arquitectura es extraer la capacidad de cómputo de rutas de los nodos que actualmente hacen esta labor y, de esta forma, éstos pueden ser más simples y baratos. Simultáneamente, en la red se situarán unos pocos nodos dedicados llamados PCE que serán los encargados del cómputo de rutas para aquellos nodos que lo requieran. Estos elementos PCE pueden ser dotados con capacidades de encaminamiento avanzadas que tengan en cuenta restricciones de ingeniería de tráfico y que satisfarán todas las necesidades de cómputo de los nodos clientes.

Este trabajo está financiado, en parte, por la Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura (España) a través del proyecto AGILA-2, con código PR1A06145, por el Ministerio de Industria, Turismo y Comercio español mediante el proyecto MESEAS, con código FIT-350301-2007-14 y por el Ministerio de Educación y Ciencia español mediante el proyecto CEPS, con código TSI2005-07520-C03-02.

Un PCE es capaz de computar LSP (*Label Switched Paths*) sobre su propio dominio porque tiene visibilidad topológica. Sin embargo, cuando la ruta se extiende a más de un dominio, se ve obligado a cooperar con PCE de esos otros dominios, que computarán segmentos de la ruta total que finalmente serán ensamblados y devueltos al PCE inicial. En esta situación, el elemento PCE puede conocer varios PCE candidatos. El mecanismo utilizado para elegir a uno de ellos tiene un impacto crucial en el tiempo que se necesitará para computar la ruta interdominio completa. Sin embargo, no existe aún un mecanismo completamente aceptado para llevar a cabo la selección de elementos PCE en entornos interdominio y, por tanto, este tema es aún objeto de estudio [7].

B. Aportaciones y estructura de este trabajo

La intención general de nuestro trabajo es contribuir al desarrollo de la arquitectura PCE. Para ello, aportamos un mecanismo para seleccionar certeramente un PCE colaborador teniendo en cuenta el estado de la red. Este mecanismo se basa en el intercambio seguro de información de TE en entornos interdominio, razón por la cual, además, aportamos un detallado estudio sobre la información de ingeniería de tráfico proporcionada por los IGP más extendidos.

El resto de este trabajo está organizado como sigue: en la segunda sección, presentamos una breve introducción a la arquitectura PCE. En la tercera sección presentamos nuestra propuesta para la selección de elementos PCE en entornos interdominio. En el cuarto punto, se muestran los aspectos claves de nuestro mecanismo. En la sección quinta presentamos nuestras conclusiones y el trabajo futuro.

II. ARQUITECTURA PCE: BREVE DESCRIPCIÓN

A. Arquitectura PCE básica

La arquitectura PCE está siendo desarrollada actualmente; por eso la mayoría de RFC (*Request For Comments*) publicados por el IETF son definiciones y requisitos generales de la arquitectura. En su configuración más básica debe contar, al menos, con tres elementos clave (Fig 1). El PCE es el nodo encargado de computar rutas. El PCC (*Path Computation Client*) es el elemento que solicitará al PCE el cómputo de rutas; y PCEP (*Path Computation Element communication Protocol*) [8], [9], [10], el protocolo de comunicaciones a través del cual se comunicarán el PCE y el PCC. Aunque a primera vista puede parecer un modelo muy simple, existen ciertas dificultades a la hora de integrar la arquitectura PCE en un dominio con las tecnologías existentes. Por ejemplo, la relación entre PCE, IGP y EGP (*Exterior Gateway Protocol*) o el modo en el que los protocolos existentes surten al PCE de información de TE.

En la arquitectura PCE, un dominio MPLS puede contar con varios PCE encargados de computar rutas sobre él. Cada nodo que quiera iniciar el establecimiento de LSP debe actuar como PCC, así que al menos los nodos LER deben actuar como PCC porque ellos son quienes establecen LSP en el interior del dominio. Además, otros nodos intermedios pueden requerir actuar como PCC si están involucrados en mecanismos de restauración local de LSP.

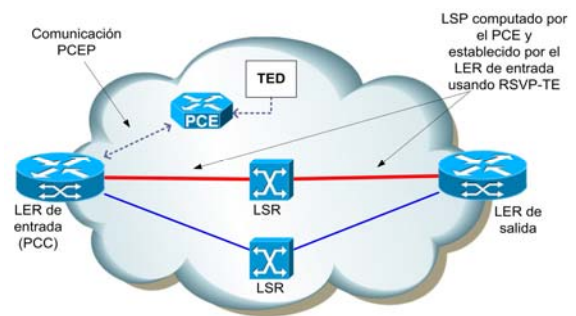


Figura 1. Componentes claves de una arquitectura PCE básica.

Cuando un flujo llega al LER de entrada al dominio, este actuará de PCC y solicitará a su PCE el cómputo de un LSP desde él hasta el destino. Para ello, utilizará el protocolo PCEP que proporciona suficiente funcionalidad para permitir realizar esta petición de forma muy flexible. La petición incorporará un conjunto de restricciones que deben ser tenidas en cuenta por el PCE a la hora de computar la ruta. El PCE computará la ruta basándose para ello en la información contenida en su TED (*Traffic Engineering Database*), que es una base de datos que incluye un grafo de estado de enlace y cualquier otra información útil. Cada PCE tiene asociada una TED que es actualizada periódicamente por los IGP y por otros mecanismos que puedan ser definidos. Una vez que el PCE ha calculado la ruta, utilizará de nuevo PCEP para devolver el resultado al LER/PCC que realizó la solicitud.

En la descripción de requisitos del protocolo PCEP se especifica que una ruta calculada, incluida en la respuesta a un LER/PCC, debe ser directamente transformable en un objeto ERO [11] (*Explicit Routing Object*) de RSVP-TE (*Resource Reservation Protocol – Traffic Engineering*); de esta forma, el LER/PCC puede iniciar el establecimiento del LSP utilizando RSVP-TE y dicho objeto ERO.

Este es el modo de funcionamiento más básico (*simple path computation*) de la arquitectura PCE, pero se permite la existencia de situaciones más complejas de gestionar y coordinar. Por ejemplo, puede existir más de un elemento PCE en el dominio, cada uno de ellos encargado de calcular LSP completos sobre él. En este caso un LER/PCC tendrá la posibilidad de elegir aquel que se ajuste más a sus necesidades. Esta situación obliga al LER/PCC a conocer las características de cada PCE para poder realizar su elección con un criterio razonable. Otro ejemplo es aquel en el que hay más de un elemento PCE en un dominio pero cada uno de ellos está encargado de calcular segmentos de LSP relativos a un área concreta de la red (*multiple path computation*). En este caso, los PCE tendrán la obligación de cooperar, computar cada uno un segmento específico, ensamblarlos y devolver al LER/PCC la ruta que solicitó. Por esta razón, hay veces en que un PCE puede actuar como PCC de cara a otros PCE.

B. Arquitectura PCE interdominio.

La situación más compleja tiene lugar cuando todo lo anterior coincide en el tiempo y, además, la ruta que se desea calcular se extiende más allá del propio dominio local. En este caso, la arquitectura PCE debe contar con mecanismos para

abordar los problemas tradicionales del encaminamiento interdominio, la ingeniería de tráfico y MPLS [12], [13]: visión parcial de la topología, falta de información de ingeniería de tráfico, seguridad, recuperación de la red o confidencialidad de la información interna de los dominios.

El funcionamiento de la arquitectura PCE en entornos interdominio sigue una serie de pasos bien estructurados dentro de los cuales es necesaria la cooperación [14], [15]; es bastante similar al funcionamiento de PCE en entornos interiores. Primero, un PCE debe conocer la existencia de elementos PCE en los dominios circundantes. Existen diversas propuestas de mecanismos de descubrimiento de elementos PCE en interdominio; por ejemplo, las definidas en [16], [17], [18], [19] o [20], todas ellas cumpliendo los requisitos expresados en [21] y la mayoría siguiendo principios similares (Fig. 2).

Una vez que el PCE ha descubierto otros elementos PCE en los dominios adyacentes, contará con una base de datos de candidatos a colaborar a la hora de recibir una petición de cómputo de LSP interdominio. En dicho momento, deberá llevar a cabo un proceso de selección del PCE adecuado. Para ello, el PCE inicial deberá contar con información relativa a cada uno de los PCE candidatos, información que es proporcionada habitualmente por los mecanismos de descubrimiento interdominio junto con el anuncio del PCE. Cuando el proceso finaliza, el PCE inicial se comunicará con el PCE seleccionado utilizando el protocolo PCEP. El resto de la operación se lleva a cabo de forma similar al trabajo intradominio excepto por el hecho de que el PCE inicial debe confiar en que el PCE colaborador computará el segmento de LSP que le corresponde cumpliendo las mismas restricciones (no tendrá visibilidad de los dominios adyacentes). Este proceso se repetirá hacia delante, dominio por dominio, hasta que se alcance el dominio final o destino del LSP.

C. Selección de PCE. Tema abierto en la arquitectura PCE

Existen todavía aspectos clave que resolver antes de contar con una arquitectura PCE interdominio funcional. Aunque se están realizando muchos esfuerzos al respecto, no hay aún un mecanismo de descubrimiento de PCE en interdominio comúnmente aceptado; y en relación con el mecanismo de selección de elementos PCE en entornos interdominio, la situación es aún peor. En los siguientes párrafos seguiremos un ejemplo explicativo para entender las dificultades y el impacto de seleccionar el PCE correcto en entornos interdominios. En la Fig. 3 podemos observar un sistema interdominio donde un PCE del dominio local tiene que computar un LSP hacia el dominio destino.

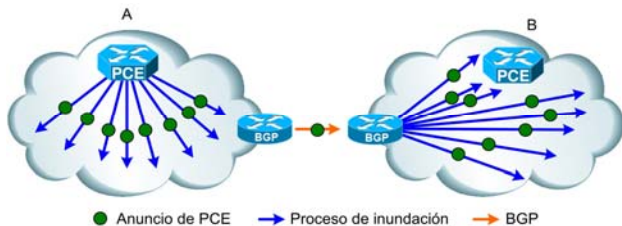


Figura 2. Mecanismo genérico de descubrimiento de PCE en interdominio.

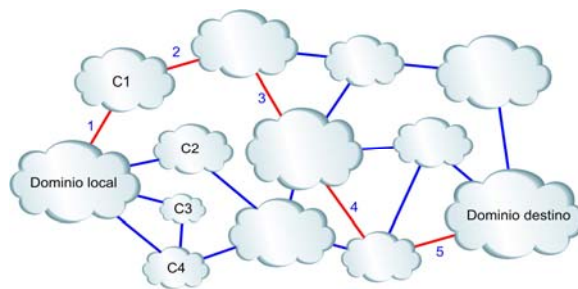


Figura 3. La mejor opción la ofrece el PCE de C1.

Como tiene cuatro dominios adyacentes (C1-C4), descubrirá PCE candidatos a colaborar en todos ellos. Hay diferencias significativas a la hora de elegir elementos PCE de cada uno de ellos como colaboradores y podría tener un gran impacto en el tiempo total necesario para computar el LSP interdominio completo. La figura muestra la mejor opción; eligiendo el PCE de C1, serán necesarias cinco colaboraciones entre PCE adyacentes para computar el LSP deseado.

Cada cooperación entre PCE implica un tiempo T_{PCEP} que incluye el tiempo utilizado por cada PCE para computar su segmento y el tiempo requerido por el protocolo PCEP (previsiblemente mayor) para comunicar a ambos PCE. T_{PCEP} es proporcional al número de dominios que atravesará el LSP.

En el ejemplo, en el mejor caso el tiempo necesario para computar el LSP es de $5 \cdot T_{PCEP}$. La Fig. 4 muestra el peor caso. Aquí, el LSP calculado es el mismo que en la Fig. 3, pero han sido necesarios varios intentos fallidos para conseguir el cálculo correcto. En total, han sido necesarias diez cooperaciones por lo que el tiempo empleado es $10 \cdot T_{PCEP}$. En general, el proceso de computar LSP interdominio sigue un proceso de *backtracking* que necesita un tiempo de $n \cdot T_{PCEP}$. Por tanto, para decrementar el tiempo necesario para computar un LSP interdominio es necesario minimizar n (el número de cooperaciones entre PCE). Pero esta no es una tarea sencilla debido a la visión parcial de la topología interdominio que tiene el PCE del dominio local. Por eso las propuestas actuales de mecanismos de selección de PCE en entornos interdominio [7] no se basan en el estado de la red. Solo tienen en cuenta el estado (o el estado inferido) de los potenciales PCE colaboradores; y precisamente esto es lo que nos proponemos mejorar.

III. PROPUESTA DE MECANISMO DE SELECCIÓN DE PCE EN ENTORNOS INTERDOMINIO

Uno de los principales problemas en la selección del PCE colaborador en entornos interdominio es la falta de visibilidad que tiene el PCE iniciador. Como hemos visto en el ejemplo, el tiempo total necesario para computar la ruta varía dependiendo del número de intentos fallidos. Estos intentos fallidos son en la mayoría de las situaciones, los responsables del incremento en el tiempo total de la computación. Para solucionar la raíz del problema, hemos diseñado un mecanismo que dote al elemento PCE de suficiente visibilidad (en términos de TE) para seleccionar de forma correcta el PCE con el que colaborará, teniendo en cuenta el estado de la red en los dominios circundantes (y hacia el dominio destino) y no sólo el estado de los propios elementos PCE.

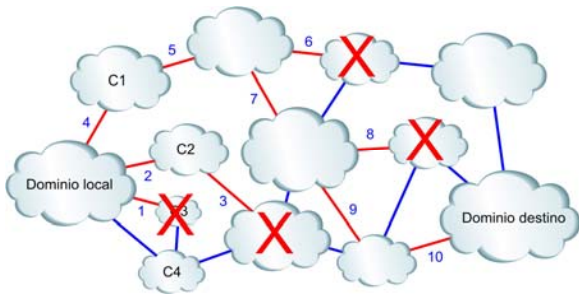


Figura 4. La peor opción la ofrece el PCE de C3.

A. Visión general de la solución

Podemos resumir nuestra propuesta como un mecanismo que exporta periódicamente información de TE entre PCE de dominios adyacentes que lo hayan acordado. La información compartida por los PCE no será sólo la relativa al dominio de cada PCE, sino a todos aquellos dominios que éstos conocen.

En relación a este método, surgen al menos dos problemas:

- Compartir información de TE en entornos interdominio podría revelar datos sensibles para la confidencialidad de los ISP, lo cual no es conveniente.
- Cada dominio puede estar utilizando un IGP diferente y los parámetros de TE proporcionados por ellos podrían ser incompatibles e incomparables entre sí.

Por ello, hemos enfocado nuestro trabajo a solventar estos problemas. Primero, hemos realizado un profundo estudio de la información de TE proporcionada por OSPF-TE (Open Shortest Path First – Traffic Engineering) [22] e ISIS-TE (Intermediate System to Intermediate System – Traffic Engineering) [23] para explorar sus diferencias y similitudes. Después, hemos seleccionado aquellos más importantes, comunes y comparables para contar con una base homogénea que permita el desarrollo de propuestas independientes del IGP. En segundo lugar, aportamos un mecanismo para agrupar toda la información de TE concerniente a un dominio dado, resultando en una visión simplificada de los recursos de dicho dominio (Fig. 5). Luego, esta información es propagada a los dominios adyacentes, donde será agregada gracias a un conjunto de funciones de agregación que hemos diseñado.

La idea es que un PCE tenga información que represente no solo el estado de un dominio adyacente, sino el estado de la ruta desde el dominio local hasta el dominio destino, a través de cualquiera de los dominios adyacentes (Fig. 6). Cuando la información proveniente de diversos dominios adyacentes es agregada, el PCE que la recibe no tiene la capacidad para distinguir la aportación individual del resto de PCE a dicha información agregada.

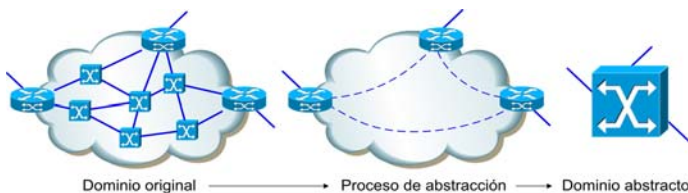


Figura 5. Proceso de abstracción de los recursos de un dominio.



Figura 6. Propagación de parámetros de TE abstractos.

De esta forma se asegura la confidencialidad de la infraestructura de los dominios a la vez que la información de TE resultante sigue siendo útil, ya que cuando un PCE tiene esta información, puede realizar sobre ella un preprocesado de las restricciones para estimar el siguiente dominio (un PCE de él) con el que colaborar. El resultado de esta estimación heurística es que el PCE seleccionado ofrecerá la menor tasa de intentos fallidos, minimizando n y decrementando de este modo el tiempo total de cómputo del LSP interdominio.

B. Clasificación y comparación de los parámetros de TE manejados por los IGP

La definición de los parámetros TE proporcionados por los IGP se encuentran muy repartidos por un gran número de RFC. Este hecho hace difícil para los investigadores encontrar la definición de un parámetro concreto cuando los necesitan. Hemos basado nuestro estudio en los RFC más importantes [22], [23], [24], [25], [26], [27], [28], [29] que definen parámetros de TE o información relativa TE para OSPF-TE e ISIS-TE. Esta información está agrupada, generalmente, en tres categorías diferentes: información relativa a los enlaces, encaminadores y a los circuitos virtuales precalculados.

En las tablas 1, 2 y 3 podemos ver el resultado de nuestro estudio relativo las tres categorías citadas. Cada tabla tiene cinco columnas: la primera describe el parámetro de TE; la segunda muestra el RFC donde está definido ese parámetro (si procede) para OSPF-TE; la tercera es similar a la segunda pero para ISIS-TE. Las columnas cuatro y cinco nos dicen si el parámetro tiene igual formato y significado para ambos IGP (✓) o no (✗).

Una de las primeras conclusiones que se obtienen de la información presentada es que OSPF-TE es más avanzado que ISIS-TE en cuanto a la información de TE que puede manejar. Sólo el 35% de los parámetros de TE que hemos estudiados están definidos para OSPF-TE e ISIS-TE a la vez. OSPF-TE tiene el 60% del total de parámetros existentes pero dichos parámetros solo están definidos para él. Por último, solo el 5% de los parámetros estudiados están definidos exclusivamente para ISIS-TE.

Debido a que nuestra intención es encontrar una base común para desarrollar nuestra estrategia, nos interesa sólo la información compartida por ambos IGP. Pero no toda ella, sino aquella que además sea común, comparable, con el mismo significado y formato. De esta forma, podremos asegurar que los parámetros de TE son entendidos correctamente y de la misma forma en todos los dominios que componen el sistema interdominio, al margen del IGP que cada uno esté utilizando. Se puede ver este subconjunto de parámetros en la tabla 4.

TABLA I. COMPARATIVA DE INF. DE TE RELATIVA A ENLACES

Información sobre el enlace	OSPF-TE	ISIS-TE	Format. similar	Signific. similar
Link type	RFC3630	-	x	✓
Traffic engineering metric	RFC3630	RFC3784	x	x
Administrative group / Colour	RFC3630	RFC3784	✓	x
Maximum bandwidth	RFC3630	RFC3784	✓	✓
Maximum reservable bandwidth	RFC3630	RFC3784	✓	✓
Unreserved bandwidth	RFC3630	RFC3784	✓	✓
TE traffic is permitted on this link	RFC4973	-	x	x
Non-TE traffic is permitted on this link	RFC4973	-	x	x
Can process IP packets	RFC4973	-	x	x
Database sync. is permitted on this link	RFC4973	-	x	x
Shared Link Risk Group (SLRG)	RFC4203	RFC4205	✓	x
Link usage cost metric	RFC4973	-	x	x
TE link maximum bandwidth	RFC4973	-	x	x
Maximum bandwidth available for TE use	RFC4973	-	x	x
Maximum available bandwidth for the LSP	RFC4203	RFC4205	✓	✓
Reserved bandwidth for TE traffic use	RFC4973	-	x	x
Link colour	RFC4973	-	x	x
Link type for non-packets networks	RFC4973	-	x	x
Local protection available	-	RFC5029	x	x
Link excluded from local protection path	-	RFC5029	x	x
Local link protection type	RFC4203	RFC4205	✓	✓
Minimum bandwidth for the LSP	RFC4203	RFC4205	✓	✓
Maximum transfer unit	RFC4203	RFC4205	✓	✓

TABLA II. COMPARATIVA DE INF. DE TE RELATIVA A ENCAMINADORES

Información sobre el encaminador	OSPF-TE	ISIS-TE	Format. similar	Signific. similar
Can act as brunch in a P2MP LSP	RFC5073	RFC5073	✓	✓
Can act as bud in a P2MP LSP	RFC5073	RFC5073	✓	✓
Supports MPLS-TE	RFC5073	RFC5073	✓	✓
Supports GMPLS	RFC5073	RFC5073	✓	✓
Can signal P2MP LSP for MPLS-TE	RFC5073	RFC5073	✓	✓
Supports OSPF graceful restart	RFC4970	-	x	x
Can act as OSPF graceful restart helper	RFC4970	-	x	x
Can act as OSPF stub router	RFC4970	-	x	x
OSPF-TE support	RFC4970	-	x	x
Supports P2LAN	RFC4970	-	x	x
Can act as LSR	RFC4973	-	x	x
Can act as LER	RFC4973	-	x	x
Can switch packets	RFC4973	-	x	x
Can switch MPLS	RFC4973	-	x	x
Support MPLS	RFC4973	-	x	x
Supports CSPF	RFC4973	-	x	x
Switching capability	RFC3203	RFC4205	✓	✓

TABLA III. COMPARATIVA DE INF. DE TE SOBRE CIRCUITOS VIRTUALES

Información sobre el circuito virtual	OSPF-TE	ISIS-TE	Format. similar	Signific. similar
Circuit duration	RFC4973	-	x	x
Circuit setup time	RFC4973	-	x	x
Circuit teardown time	RFC4973	-	x	x
No. of TE circuit paths	RFC4973	-	x	x

TABLA IV. PARÁMETROS DE TE COMUNES Y COMPARABLES MANEJADOS SIMULTÁNEAMENTE POR OSPF-TE E ISIS-TE

Parámetro de TE o información relativa a TE	OSPF-TE	ISIS-TE	Formato y significado similares
Maximum bandwidth	RFC3630	RFC3784	✓
Maximum reservable bandwidth	RFC3630	RFC3784	✓
Unreserved bandwidth	RFC3630	RFC3784	✓
Maximum available bandwidth for the LSP	RFC4203	RFC4205	✓
Local link protection type	RFC4203	RFC4205	✓
Minimum bandwidth for the LSP	RFC4203	RFC4205	✓
Maximum transfer unit	RFC4203	RFC4205	✓
Can act as brunch in a P2MP LSP	RFC5073	RFC5073	✓
Can act as bud in a P2MP LSP	RFC5073	RFC5073	✓
Supports MPLS-TE	RFC5073	RFC5073	✓
Supports GMPLS	RFC5073	RFC5073	✓
Can signal P2MP LSP for MPLS-TE	RFC5073	RFC5073	✓
Switching capability	RFC4203	RFC4205	✓

Nuestra intención es minimizar, en la medida de lo posible, el trasiego de información entre dominios. Por ello, basándonos en los parámetros más utilizados en las diversas propuestas que existen en la literatura actual, hemos reducido la tabla 4. El subconjunto final de parámetros de TE y sus respectivos formatos se puede consultar en la tabla 5. Nuestra solución utilizará esos diez parámetros de ingeniería de tráfico que cubren un alto porcentaje de los que son comunes a OSPF-TE e ISIS-TE; simultáneamente, este subconjunto es útil para la mayoría de las propuestas actuales sobre ingeniería de tráfico en entornos interdominio.

C. Intercambio seguro de parámetros de TE en interdominio

Cada dominio abstraerá la información de sus propios recursos para anunciarse como una caja negra al resto de dominios. Si la información a mostrar fuese relativa a un solo dominio, sería muy sencillo para un PCE vecino descubrir más detalles de los deseables sobre los recursos de la red. Por ello, la abstracción de los datos es solo el primero de dos pasos; hace falta también un mecanismo que permita que la información anunciada no revele detalles concretos de las topologías y pueda ser transmitida de forma segura.

Para ello, cada dominio anunciará no exactamente la información relativa a sus propios recursos sino una agregación de dicha información y la proveniente de los dominios circundantes. Así, el dominio que reciba la información agregada puede utilizarla para seleccionar el elemento PCE con el que colaborar pero dicha información sólo le proporciona una visión difusa del sistema interdominio. La agregación hace que la información sea menos exacta, por tanto, este es un mecanismo para construir información de TE de grano grueso.

TABLA V. CONJUNTO FINAL DE PARÁMETROS DE TE IDÉNTICOS

Descripción del parámetro de TE	Nº	Formato
Maximum bandwidth	1	Formato de punto flotante de 32 bits IEEE 754.
Maximum reservable bandwidth	2	Formato de punto flotante de 32 bits IEEE 754.
Unreserved bandwidth	3	8x32 bits en formato de punto flotante de 32 bits IEEE 754. Uno para cada uno de los 8 niveles de prioridad.
Maximum available bandwidth for the LSP	4	8x32 bits en formato de punto flotante de 32 bits IEEE 754. Uno para cada uno de los 8 niveles de prioridad.
Local link protection type	5	1 octeto. Valores 0x01, 0x02, 0x04, 0x08, 0x10 y 0x20, dependiendo de la protección local seleccionada.
Minimum bandwidth for the LSP	6	Formato de punto flotante de 32 bits IEEE 754.
Maximum transfer unit	7	Un número de 2 octetos [0-2 ¹⁶].
Supports MPLS-TE	8	1 bit. Significado booleano.
Supports GMPLS	9	1 bit. Significado booleano.
Switching capability	10	1 octeto. Valores 1, 2, 3, 4, 51, 100, 151 y 200, dependiendo de la capacidad de conmutación indicada.

Hemos diseñado un conjunto de funciones de agregación que cumplen tienen unas características concretas: proporcionan información útil, ocultan los detalles sobre la contribución de cada dominio, son simples y no consumen mucho procesador, para no congestionar al PCE. Por esta razón, hemos decidido utilizar las que se muestran en la tabla 6. Cada función de agregación utiliza dos argumentos de entrada:

- El mejor valor del parámetro de TE dentro del dominio local, desde un punto de entrada hasta uno de salida del dominio para alcanzar el dominio destino (es el resultado del proceso de abstracción).
- El mejor valor anunciado por los dominios adyacentes, relativo al mismo parámetro de TE, para alcanzar al dominio destino.

El resultado de las funciones de agregación es el mejor valor que puede ser asegurado a lo largo de toda la ruta desde el dominio anunciador hasta el dominio destino. Se utiliza solo para ser anunciado a los dominios circundantes. El dominio local utilizará en su lugar la información agregada que los dominios adyacentes le anuncien a él. Es decir, un PCE utilizará la información anunciada por los dominios adyacentes para seleccionar el PCE adecuado en cada momento y a esta información añadirá su mejor visión hacia el destino (reagregará información) para anunciársela a su vez a otros.

TABLA VI. FUNCIONES DE AGREGACIÓN PARA CADA PARÁMETRO

Descripción del parámetro de TE	Nº de parámetro	Función de agreg.
Maximum bandwidth	1	Min()
Maximum reservable bandwidth	2	Min()
Unreserved bandwidth	3	Min()
Maximum available bandwidth for the LSP	4	Min()
Local link protection type	5	Min()
Minimum bandwidth for the LSP	6	Max()
Maximum transfer unit	7	Min()
Supports MPLS-TE	8	And()
Supports GMPLS	9	And()
Switching capability	10	And()

Hemos seleccionado sólo tres funciones para agregar todos los parámetros de TE. La función *Min()*, que devuelve el menor de los dos valores que recibe como parámetros, se utiliza para aquellos parámetros que son números, excepto para el mínimo ancho de banda para el LSP, que requiere el uso de la función *Max()*. La función *Max()* devuelve el mayor valor de los dos que recibe como argumentos de entrada. Finalmente, hemos elegido la función *And()* para aquellos parámetros de TE que son valores booleanos o máscaras de bits interpretadas como valores lógicos. La función *And()* realiza un AND lógico entre los dos argumentos que toma como entrada. El resultado de aplicar las funciones de agregación es transmitido al resto de dominios adyacentes afectados (a sus PCE) como una tema: el valor agregado para el parámetro, el identificador del dominio que realiza el anuncio y el identificador del dominio al que se refiere la información agregada (el dominio destino). Esto permite saber el modo en que será almacenada en la TED del PCE que la reciba: una matriz tridimensional (Fig. 7). Debido a las similitudes de esta estructura de datos con un cubo, hemos llamado a esta estructura RI-CUBE (*Routing Information - CUBE*). Podemos calcular el tamaño total de la información y el modo en que esta escala. El tamaño del RI-CUBE (Γ) sigue una fórmula (1) donde influyen el número PCE colaboradores (x), el número de dominios involucrados (y) y el número de parámetros de TE intercambiados (z). También el espacio necesario para albergar a cada parámetros de TE (φ_{ijk}).

$$\Gamma = \sum_{i=1}^x \sum_{j=1}^y \sum_{k=1}^z \varphi_{ijk} \quad (1)$$

La Fig. 8 muestra el tamaño total del RI-CUBE cuando el número de PCE que colaboran y el número total de dominios que componen el sistema interdominio crecen de cero a cien. Podemos observar que si el PCE local tiene conexión total con, hipotéticamente, 100 dominios, el tamaño total del RI-CUBE es menor de 800 KB. En cualquier caso, no se espera que la arquitectura PCE funcione con modelos tan grandes. En [8] se definen los requisitos de escalabilidad de la arquitectura y se apunta un número total de 20 dominios. Usando este número y suponiendo el peor de los casos (topología *full-connect* e intercambio de todos los parámetros de TE), el tamaño total del RI-CUBE es de 31 KB.

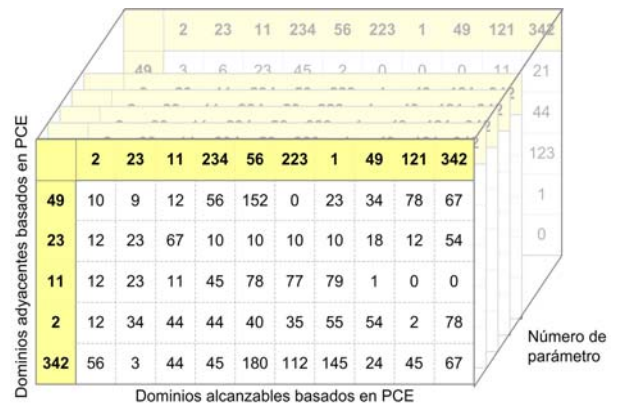


Figura 7. Almacenamiento de información de TE agregada en el RI-CUBE

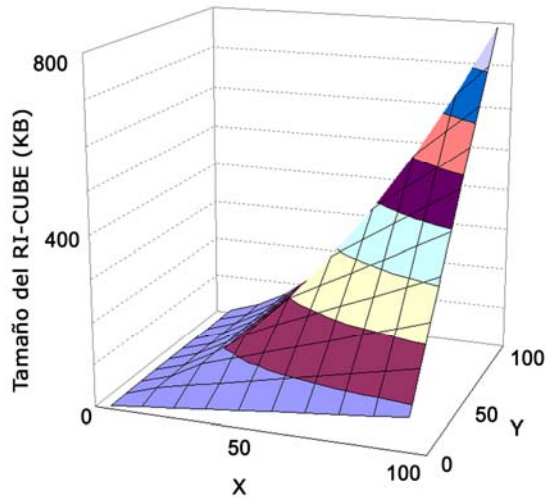


Figura 8. Estimación del tamaño del RI-CUBE y su evolución.

D. Preprocesado de restricciones

Un PCE encargado de calcular rutas interdominio que implemente esta propuesta (un PCE mejorado) debe tener un módulo para gestionar el RI-CUBE (Fig. 9). Este módulo abstrae la información de los recursos de TE en el dominio local (de la información contenida en su TED) y agrega y propaga la información de TE a PCE de otros dominios cuando se necesite. Cuando un PCE mejorado recibe una petición de cómputo de ruta interdominio de un PCC, debe seleccionar el correspondiente PCE de un dominio adyacente para colaborar. Para ello cuenta con su base de datos de PCE candidatos. Además, ahora, tiene una estructura de datos, el RI-CUBE que permite ser consultado de forma muy flexible (Fig. 10) para realizar una selección de PCE basado en el estado de la red.

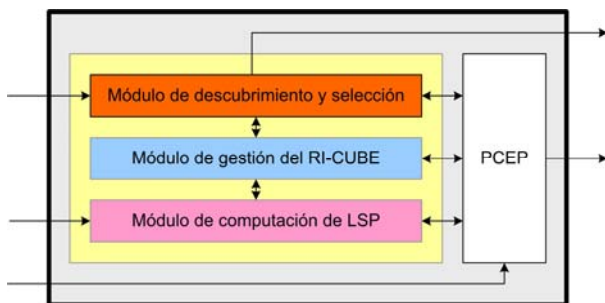


Figura 9. Arquitectura modular de un PCE mejorado.

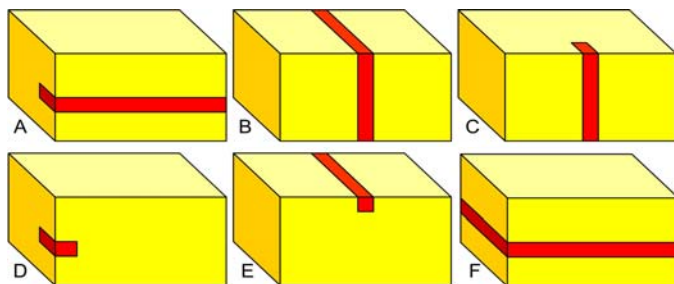


Figura 10. Tipos de consulta soportadas por el RI-CUBE

Como ejemplo, Fig. 10 A muestra la selección de un parámetro de TE para todos los dominios alcanzables desde el dominio local, mientras que Fig. 10 B representa la selección de todos los parámetros para un dominio alcanzable dado. En general, permite conocer el estado de la red, en relación a uno o varios parámetros, para llegar a dominios alcanzables a través de los distintos dominios adyacentes. Con el RI-CUBE actualizado y haciendo uso de estas consultas, un PCE mejorado puede estimar si seleccionar a un PCE como colaborador permitirá que el LSP pueda ser calculado o no. Además, puede predecir no solo la zona de la red con más recursos disponibles, sino aquella cuyos recursos se ajustan mejor a los requeridos en la solicitud de cómputo de LSP. Esto permite distribuir los recursos de la red equitativamente.

Finalmente, una vez que el dominio adyacente ha sido seleccionado, el PCE iniciador debe buscar en su base de datos de PCE candidatos para encontrar uno con capacidad para interdominio y perteneciente a dicho dominio.

IV. ASPECTOS RELEVANTES DE ESTA SOLUCIÓN

Aplicar esta propuesta en un sistema interdominio basado en PCE puede tener aspectos beneficiosos. En primer lugar, la cantidad de información en la que se basa el proceso de selección de PCE es mayor. Un PCE tradicional solo conocerá las capacidades de los PCE circundantes, especificadas en [30], [31] y [32]. Un PCE mejorado tendrá además la información de TE recogida en la tabla 5, que hace referencia al estado de la red en la trayectoria hacia el destino. En segundo lugar, como esta propuesta está pensada para mejorar el proceso en sistemas interdominio, no es necesario que todos los PCE la implementen. Sólo aquellos PCE encargados de tareas en interdominio. La arquitectura PCE se suele organizar de forma jerárquica. La raíz de esta jerarquía serán los PCE favorecidos por el uso de la información de TE agregada aquí descrita (Fig. 11). Por esta razón, el impacto en la organización interna del dominio es mínimo. En último lugar, esta técnica puede ser implementada incrementalmente, desde el núcleo del sistema interdominio hasta la periferia. La información agregada la utiliza el PCE independientemente de que esta información sea o no trasladada a terceros dominios. Por tanto, cada PCE puede aprovecharse de un PCE que utilice la información comentada, inconscientemente, al menos en alguno de los tramos del LSP. Por tanto, cuantos más dominios tengan PCE con esta propuesta, mejor será el proceso de cómputo de rutas global; el sistema interdominio completo se beneficiará de ello.

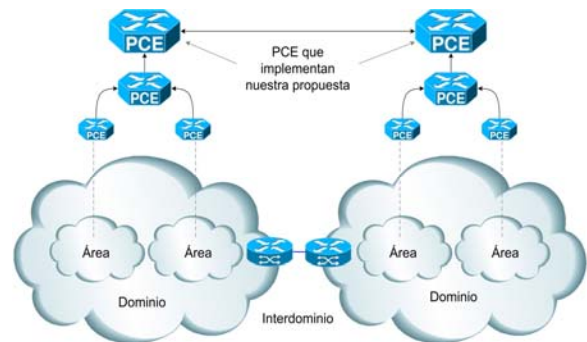


Figura 11. Configuración jerárquica de la arquitectura PCE interdominio

V. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo en curso presentamos nuestro diseño de un mecanismo para seleccionar PCE externos como colaboradores en el cálculo de una ruta interdominio, teniendo en cuenta el estado de la red hacia el dominio destino y no sólo el estado de los propios PCE. Este mecanismo se basa en la información de TE proporcionada por los IGP involucrados en el funcionamiento global del sistema interdominio. Debido a las diferencias en la información proporcionada por OSPF-TE e ISIS-TE, los IGP más importantes, hemos proporcionado un profundo estudio comparativo sobre la información de TE manejada por ambos. También hemos proporcionado un mecanismo para compartir dicha información de forma segura entre dominios mediante la agregación de parámetros de TE importantes, comparables y comunes. Con esta información en su TED, un PCE mejorado con puede preprocesar las restricciones que acompañan a una solicitud de cómputo de LSP y de esta forma seleccionar el PCE colaborador de una forma más exacta. Esto es así porque el mecanismo permite prever las expectativas de cubrir los requisitos impuestos a la ruta. Estas técnicas están justificadas debido a la importancia que tiene el proceso de selección de PCE colaboradores en el tiempo total necesario para computar un LSP interdominio.

Próximamente, trabajaremos en esta propuesta diseñando mecanismos de bajo impacto para transportar la información agregada entre dominios adyacentes. Además, profundizaremos en el diseño de algoritmos de preprocesado de restricciones, rápidos y útiles, que permitan explotar la información de TE interdominio disponibles gracias a esta propuesta.

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] A. Farrel, J. P. Vasseur, J. Ash. "A Path Computation Element (PCE)-Based Architecture". IETF RFC 4655. August, 2006.
- [4] E. Rosen, A. Viswanathan, R. Callon. "Multiprotocol Label Switching Architecture". IETF RFC 3031. January, 2001.
- [5] E. Mannie. "Generalized Multi-Protocol Label Switching (GMPLS) Architecture". IETF RFC 3945. October, 2004.
- [6] E. Oki, I. Inoue, K. Shimoto. "Path computation element (PCE)-based traffic engineering in MPLS and GMPLS networks". IEEE Sarnoff Symposium 2007. April 30 2007-May 2 2007. pp. 1-5. Digital Object Identifier: 10.1109/SARNOF.2007.4567400.
- [7] T. Saad, J. Israr, S. Sivabalan, H.T. Mouftah. "An Evaluation for PCE Selection Schemes for Inter-Domain Path Computation". 9th International Conference on Transparent Optical Networks, 2007. ICTON '07. Volume 3, 1-5. July 2007 pp. 187-187. DOI 10.1109/ICTON.2007.4296276.
- [8] J. Ash, J.L. Le Roux. "Path Computation Element (PCE) Communication Protocol Generic Requirements". IETF RFC 4657. September, 2006.
- [9] J.L. Le Roux. "Path Computation Element Communication Protocol (PCECP) Specific Requirements for Inter-Area MPLS and GMPLS Traffic Engineering". IETF RFC 4927. June, 2007.
- [10] J. P. Vasseur, J. L. Le Roux. "Path Computation Element (PCE) communication Protocol (PCEP)". IETF Draft draft-ietf-pce-pcep-17.txt. Work in progress. November, 2008.
- [11] 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.
- [12] M. Yannuzzi, X. Masip-Bruin, O. Bonaventure. "Open issues in interdomain routing: a survey". IEEE Network. Volume 9, Issue 6. November-December, 2005.
- [13] 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.
- [14] R. Zhang, J.P. Vasseur. "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements". IETF RFC 4216. November, 2005.
- [15] J.L. Le Roux, J.P. Vasseur, J. Boyle. "Requirements for Inter-Area MPLS Traffic Engineering". IETF RFC 4105. June, 2005.
- [16] K. Kumaki, T. Murai. "BGP protocol extensions for Path Computation Element (PCE) Discovery in a BGP/MPLS IP-VPN". IETF Draft draft-kumaki-pce-bgp-disco-attribute-02.txt. October, 2008. Work in progress.
- [17] Vijayanand, Somen Bhattacharya, Prasanna Kumar. "BGP protocol extensions for PCE Discovery across Autonomous Systems". IETF Draft draft-vijay-somen-pce-disco-proto-bgp-04.txt. July, 2007. Expired I-D.
- [18] M. Boucadair, P. Morand. "Path Computation Service discovery via Border Gateway Protocol". May, 2005. IETF Draft draft-boucadair-pce-disco-01.txt. Expired I-D.
- [19] M. Boucadair, P. Morand. "A Solution for providing inter-AS MPLS-based QoS tunnels". IETF Draft draft-boucadair-pce-interas-01.txt. May 2005. Expired I-D.
- [20] M. Domínguez-Dorado, José-Luis González-Sánchez, J. Domingo-Pascual. "PILEP: a contribution to PCE-based interdomain path computation". Proceedings of the 13th International Telecommunications Network Strategy and Planning Symposium (NETWORKS'08). pp. 1-7. ISBN 978-963-8111-68-5. Budapest (HUNGARY), October, 2008.
- [21] J.L. Le Roux. "Requirements for Path Computation Element (PCE) Discovery". IETF RFC 4674. October, 2006.
- [22] K. Kompella, D. Yeung. "Traffic Engineering (TE) Extensions to OSPF Version 2". IETF RFC 3630. September, 2004.
- [23] H. Smit, T. Li. "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)". IETF RFC 3784. June, 2004.
- [24] P. Srisuresh, P. Joseph. "OSPF-xTE: Experimental Extension to OSPF for Traffic Engineering". IETF RFC 4973. July, 2007.
- [25] K. Kompella, Y. Rekhter. "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)". IETF RFC 4203. October, 2005.
- [26] K. Kompella, Y. Rekhter. "Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)". IETF RFC 4205. October, 2005.
- [27] J.P. Vasseur, S. Previdi. "Definition of an IS-IS Link Attribute Sub-TLV". IETF RFC 5029. September, 2007.
- [28] J.P. Vasseur, J.L. Le Roux. "IGP Routing Protocol Extensions for Discovery of Traffic Engineering Node Capabilities". IETF RFC 5073. December, 2007.
- [29] A. Lindem, N. Shen, J.P. Vasseur, R. Aggarwal, S. Shaffer. "Extensions to OSPF for Advertising Optional Router Capabilities". IETF RFC 4970. July, 2007.
- [30] 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.
- [31] 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.
- [32] M. Domínguez-Dorado, José-Luis González-Sánchez, J. Domingo-Pascual. "Descubrimiento de PCE inter-AS: una aportación a la computación de LSP en sistemas multidominio". Libro de actas de las VII Jornadas de Ingeniería Telemática (JITEL'08). Págs. 80-86. ISBN 978-84-612-5474-3. Alcalá de Henares (ESPAÑA), Septiembre de 2008.