Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

Cisco „Campus Fabric“ oder Software-defined Access

von Dr. Joachim Wetzlar

Dr.-Ing. Joachim Wetzlar ist seit mehr denn 20 Jahren Senior Consultant der ComConsult Beratung und Planung GmbH und leitet dort das Competence Center „Data Center“. Er verfügt über einen erheblichen Erfahrungsschatz im praktischen Umgang mit Netzkomponenten und Serversystemen. Seine tiefen Detailkenntnisse der Kommunikations-Protokolle und entsprechender Messtechnik haben ihn in den zurückliegenden Jahren zahlreiche komplexe Fehlersituationen erfolgreich lösen lassen. Neben seiner Tätigkeit als Troubleshooter führt Herr Dr. Wetzlar als Projektleiter und Senior Consultant regelmäßig Netzwerk- WLAN- und RZ-Redesigns durch. Besucher von Seminaren und Kongressen schätzen ihn als kompetenten und lebendigen Referenten mit hohem Praxisbezug.

Ohne das Attribut “Software-defined” lassen sich derzeit offensichtlich keine neuen Produkte mehr an den Mann bringen. Nach den Software-defined Networks (SDN) haben wir an dieser Stelle das Software-defined Data Center (SDDC) diskutiert. Sie erinnern sich, es ging um den Einsatz so genannter Overlays und um die Frage, ob man diese besser auf dem Server bzw. Hypervisor oder in den Netzkomponenten realisiert. Nun also Software-defined Access (SDA).

Eigentlich hätte mich dieses Thema nur am Rande interessiert, denn ich nehme bezüglich derlei Techniken grundsätzlich erst einmal eine abwartende Haltung ein. Letztlich hinter dem Ofen hervorgelockt hat mich die Tatsache, dass Cisco mit seiner „Campus Fabric“ eine spannende Anwendung des Locator/ID Separator Protocols (LISP) gelungen ist. Darüber hinaus erkenne ich auch grundsätzlichen Nutzen des SDA für einige meine Projekte.

Stellen Sie sich also die folgende Aufgabenstellung vor: Die Access-Netze eines Kunden basieren auf den typischen drei Stufen „Access“, „Distribution“ und „Core“. Es handelt sich um einen weit verzweigen Campus mit zahlreichen Gebäuden und entsprechend mehr als einer Handvoll Distribution-Bereichen. Insgesamt gibt es eine dreistellige Anzahl Access Switches und eine entsprechend große Anzahl von Access-Subnetzen. Ein IP-Adressierungskonzept berücksichtigt die geographische Aufteilung der Subnetze und erwartete Client-Anzahlen. Sie können sich sicher vorstellen, wie ein solches IP-Adressierungskonzept nach einigen Jahren praktischem Netzbetrieb mit „Moves, Adds und Changes“ aussieht.

Nun wird gefordert, Network Access Control (NAC) einzuführen. Clients werden verschiedenen Sicherheitszonen zugeordnet. Die Einteilung erfolgt bezogen auf Benutzergruppen, also etwa „Standard-Arbeitsplatz“, „Personalabteilung“, „Vorstand“, „Entwicklung“ und „Gäste“. Selbstverständlich werden Sie zu diesem Zweck nicht mehrere physische Netze parallel aufbauen und betreiben wollen. Stattdessen wählen Sie eine Technik zur Netzwerk-Virtualisierung, wie beispielsweise Virtual Routing and Forwarding (VRF) und Virtual LAN (VLAN). Oder sie machen es per Overlay, z.B. mittels Multi-Protocol Label Switching (MPLS).

Ein Problem bleibt jedoch bei den meisten Techniken: Sie müssen unterschiedlichen Benutzergruppen unterschiedliche IP-Subnetze zuweisen. Man kann sich die Sache einfach machen und das IP-Adressierungskonzept vervielfachen. Wo es bisher ein Access VLAN gab, sind es nun fünf. Welch ein Aufwand, vor allem weil die meisten Benutzergruppen nur wenige Anwender umfassen. Also könnte man deren Subnetze auf bestimmte Bereiche begrenzen. Aber was passiert bei Umzügen? Wählen wir also unterschiedliche geographische Ausdehnungen: Standard-Arbeitsplätze

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

bleiben wie gehabt und zusätzlich je ein Campus-übergreifendes Netz für Entwicklung, Personalabteilung, etc. Aber auch das hat wohlbekannte Nachteile.

Sie erkennen, dass eine wie auch immer geartete neue Technik nützlich wäre, Access-Netze von der örtlichen Bindung an das IP-Subnetz der Switching-Infrastruktur zu entkoppeln. Anders ausgedrückt, man braucht eine Switching-Infrastruktur, mit der sich der Zugang zu bestimmten IP-Subnetzen ortsunabhängig realisieren lässt. Dies lässt sich mit spezieller Hardware realisieren, wie beispielsweise mittels Bridge Port Extension. Access Switches mutieren dabei zu abgesetzten Line Cards eines großen modularen „Access Core". Oder man greift (mal wieder) auf ein Overlay zurück, wie beispielsweise Shortest Path Bridging MAC (SPBM).

Cisco hat nun eine neues Overlay für diesem Zweck erfunden, die „Campus Fabric". Das Overlay ist hierbei eine Kombination aus LISP und Virtual Extensible LAN (VXLAN). Schauen wir uns zunächst die Funktionsweise vom LISP an; ich vereinfache so weit wie möglich.

Grundzüge von LISP

LISP, das Locator/ID Separator Protocol, wurde im Januar 2013 als RFC 6830 veröffentlicht. Die in der Überschrift zu dem als „experimental" eingestuften Dokument genannten Autoren stammen allesamt von Cisco Systems. Bereits im Sommer 2011 hat Cisco in unserem Hause eine Implementierung auf Basis der Serie Nexus 7000 vorgeführt und wir konnten erste Tests der Funktionsweise im ComConsult-Labor durchführen. Kurz gesagt, unterscheidet LISP zwei Ebenen der IP-Adressierung:

- Ebene der Endpunkte, die miteinander kommunizieren: Jeder Endpunkt verfügt über eine Adresse, den so genannten End Point Identifier (EID). Dieser kann entweder eine IPv4 oder eine IPv6-Adresse sein. Die EIDs sind Teil lokaler IP-Netze, wie beispielsweise Access-Netze oder Rechenzentren.

- Ebene des „Underlay": Hierbei handelt es sich um ein beliebiges geroutetes IPv4- oder IPv6-Netz. Die Netze der EIDs sind daran über spezielle Router angeschlossen, die eine Enkapsulierung bzw. Dekapsulierung der Pakete von EIDs vornehmen. Aus dem Underlay werden diese Router über so genannte Routing Locators (RLOC) adressiert. Der RLOC ist also die IP-Adresse des LISP-Routers aus der Sicht des Underlay.

Die Übertragung von Paketen erfolgt im Underlay getunnelt. IP-Pakete werden in IP-Paketen verpackt, wobei alle vier Varianten denkbar sind, also IPv4 in IPv4, IPv4 in IPv6, usw. Dementsprechend heißen diese Router auch „Tunnel Router". Es werden zwei grundsätzliche Funktionen beim Tunnel Router unterschieden:

- Ingress Tunnel Router (ITR): Der ITR enkapsuliert Pakete der EIDs und sendet sie über das Underlay an einen ETR.

- Egress Tunnel Router (ETR): Der ETR nimmt enkapsulierte Pakete aus dem Underlay entgegen, dekapsuliert sie und sendet den Inhalt per „normalem" IP Routing an den entsprechenden EID.

Jeder Router unterstützt im Allgemeinen beide Funktionen, damit den Paketen ein Rückweg offensteht. Die Router werden daher gerne mit dem Kürzel xTR bezeichnet. Die Abbildung 1 zeigt ein einfaches Beispiel für LISP. Neben einem Access-Netz und dem Rechenzentrum erkennt man das Underlay und die beiden xTR mit ihren IP-Adressen (RLOC). Außerdem sind ein Client und ein Server mit ihren IP-Adressen (EID) zu erkennen.

Eine Funktion fehlt noch, damit die Pakete letztlich ihr Ziel erreichen. Es muss eine Abbildung von EID auf die RLOC geben. Der ITR muss irgendwo nachschlagen können, an welchen ETR er das enkapsulierte Paket senden soll. Dazu spezifiziert LISP die beiden Funktionen „Map Resolver" und „Map Server". Im Map Server hinterlegen die ETR alle über sie erreichbaren EID bzw. deren IP-Netze. Die entsprechende Tabelle ist in Abbildung 1 angedeutet. Man sieht dort zwei Einträge, einen für das Access-Netz und einen für das Rechenzentrum. Mit Pfeilen angedeutet ist, wie der ITR die Adresse des passenden RLOC erfährt:

1. Der Client im Access-Netz sendet sein Paket an den ITR. Die IP-Adresse des ITR auf der Seite des Access-Netzes (z.B. 192.168.2.1) könnte am Client als Default Gateway eingerichtet sein. Der ITR bittet nun den Map Resolver, den zum Ziel-EID passenden RLOC zu suchen.

2. Der Map Resolver gibt die Anfrage an den Map Server weiter.

3. Der Map Server findet den gesuchten Eintrag und leitet die Anfrage an den RLOC weiter. Dabei handelt es sich um den bzw. einen gesuchten ETR.

4. Dieser ETR informiert nun den anfragenden ITR über seine IP-Adresse im Underlay (RLOC). Der ITR legt die Adresse in seinem Cache ab. Er sendet nun das enkapsulierte Paket an den ETR.

Dieser etwas komplizierte Ablauf stellt einerseits sicher, dass der ETR tatsächlich eine Verbindung zum ITR besitzt (und nicht z.B. inzwischen ausgefallen ist). Zum anderen besteht so die Möglichkeit, abhängig von der Verkehrssituation stattdessen den RLOC eines zweiten Interfaces oder eines redundanten ETR zu bekanntzugeben und damit letztlich den Datenfluss zu steuern.

Mobilität von Endpunkten mit LISP

Sie werden bemerkt haben, dass die Frage, wo sich ein Endpunkt (also Client oder Server) befindet, unabhängig von der Netzwerk-Adresse in den jeweiligen Access-Netzen ist. Diese Übereinstimmung von EID und Netzwerk-Adresse wird nur für das Routing innerhalb der Access-Netze benötigt.

Machen wir also folgendes Gedankenspiel: Die Access-Netze umfassen jeweils nur ein /24-Subnetz und der xTR ist für dieses Subnetz der Default Gateway. Nun nehmen wir einen Client aus einem Access-Netz heraus und stecken ihn auf ein Switch Port, das sich in einem zweiten Access-Netz befindet. Die IP-Adresse des Client bleibt jedoch unverändert, es kommt also insbesondere kein DHCP zum Einsatz. Abbildung 2 illustriert dieses Szenario. Jetzt passiert folgendes:

- Der umgezogene Client versucht den Server zu erreichen und sendet ein entspre-

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

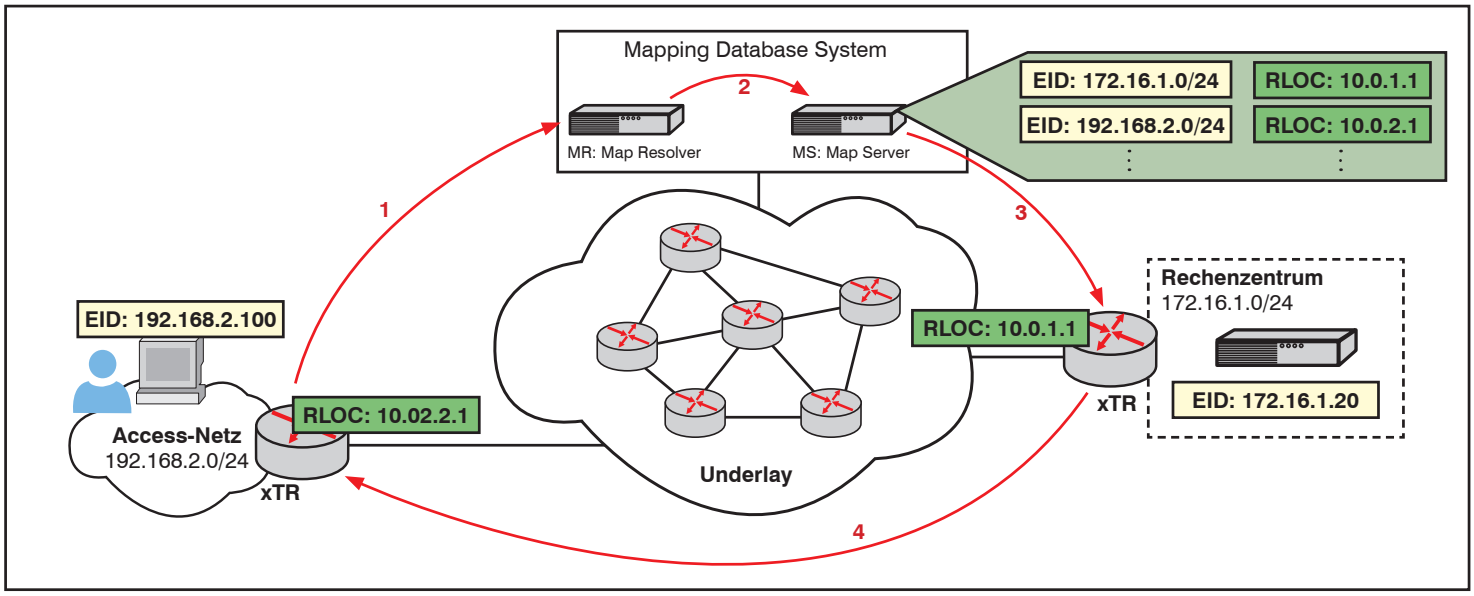

Abbildung 1: Zur grundsätzlichen Funktionsweise von LISP

chendes Paket an den ITR (wie das einzelnen geht, lassen wir hier unberücksichtigt).

- Der ITR erfährt über den Map Resolver den RLOC des passenden ETR und sendet das Paket enkapsuliert über das Underlay, wie oben beschrieben.
- Der ITR hat bei dieser Gelegenheit bemerkt, dass sich nun ein Client in seinem Access-Netz befindet, der nicht zu ihm gehört, dessen IP-Adresse aus einem anderen IP-Netz stammt.
- Nun registriert der ITR die IP-Adresse des Client am Map Server. Er verwendet dazu die /32-Subnetzmaske, die nur genau die eine IP-Adresse bezeichnet. Es entsteht ein neuer Eintrag im Map Server, den Sie in der Abbildung 2 erkennen können. Gleichzeitig erzeugt der ITR eine Host Route für den Client, die in das angeschlossene Access-Netz zeigt.
- Der Map Server informiert nun den ITR des ursprünglichen Access-Netzes (192.168.2.0/24) darüber, dass sich der Client nicht mehr in seinem Bereich befindet. Er macht sich einen entsprechenden Eintrag in einer „Abwesenheits-Liste“.

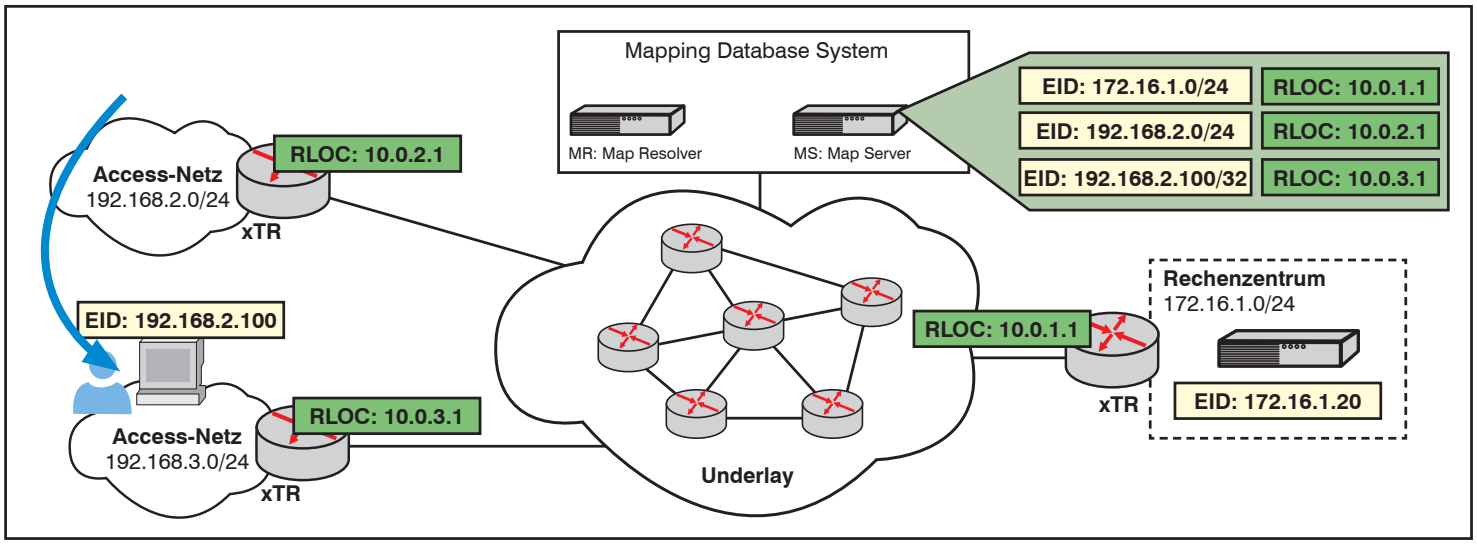

Abbildung 2: LISP unterstützt die Mobilität von Endpunkten

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

Nach diesen Maßnahmen würde das Antwortpaket des Servers zunächst am ursprünglichen RLOC ankommen. Dieser xTR führt den Client in der Abwesenheits-Liste und wird den absendendenden ITR im Rechenzentrum anweisen, seinen entsprechenden Cache-Eintrag zu löschen. Dieser führt dann die Standard-LISP-Prozedur mit Anfrage am Map Resolver durch und erhält nun als neuen RLOC die IP-Adresse 10.0.3.1 genannt.

Verwendung von LISP in Cisco Campus Fabric

Sie merken, dass die Sache mit der Mobilität von Endpunkten so nicht ganz rund ist. Es mutet irgendwie komisch an, dass sich Endgeräte in Subnetzen aufhalten sollen, in die sie eigentlich nicht hineingehören. An dieser Stelle würde man doch lieber auf DHCP vertrauen und den Endpunkten neue IP-Adressen und auch das passende Default Gateway zuweisen. In diesem Fall kann man sich fragen, wozu man dann noch den Umweg über das Underlay braucht. Ganz Genau!

Cisco hat nun einen entscheidenden Kunstgriff gewagt: Alle Access-Netze haben dieselbe IP-Adresse und Subnetzmaske. Wie bitte? Richtig: Es sieht nach einem „Split Brain“ aus. Dadurch würde eigentlich jegliche Kommunikation mit dem entsprechenden Subnetz unmöglich. Nicht so dank der Fähigkeiten von LISP. Schauen Sie sich hierzu die Abbildung 3 an.

Alle Access-Netze tragen dieselbe IP-Netzwerkadresse und Subnetzmaske. Mehr noch: Die xTR haben aus Sicht der Clients alle dieselbe IP-Adresse (z.B. 192.168.2.1) und MAC-Adresse. In dieser Hinsicht ähnelt das Verfahren also den Protokollen zur Router-Redundanz, wie dem Virtual Router Redundancy Protocol (VRRP) oder seinem Cisco-Pendant HSRP (Hot-Standby Router Protocol). Aus Sicht der Clients ist die Sache also einfach. Egal, in welchem physischen Netz sie sich befinden, das Default Gateway wird immer erreicht. Cisco bezeichnet das als „Anycast Gateway“.

Da nun alle Access-Netze dieselbe Adresse tragen, kann es für diese Netze keine Einträge mehr im Map Server geben. Es lässt sich kein eindeutiger RLOC für ein Netz bestimmen. Es gibt nur noch Host-Einträge, denn Hosts lassen sich eindeutig einem Access-Netz und seinem ETR bzw. RLOC zuordnen. Auf der Seite des Rechenzentrums könnte das grundsätzlich genauso sein, wenn der xTR unmittelbar mit dem Layer-2-Segment in Verbindung stünde, an dem sich auch die Server befinden. Wird dazwischen geroutet, erscheinen wie gehabt ganze Netze im Map Server, wie in Abbildung 3 dargestellt. Zur Unterscheidung der Funktionsweise, werden ETR bzw. ITR hier mit dem Zusatz „Proxy“ belegt, kurz also PxTR.

Wenn Sie nun die oben beschriebene Standard-LISP-Prozedur durchspielen, werden Sie nachvollziehen können, dass Clients in Access-Netzen erfolgreich mit Servern im Rechenzentrum kommunizieren können und umgekehrt. Und wir haben nun eine Lösung für das Problem des Kunden, der die 100 auf zwei Gebäude und sechs Etagen verteilten Mitarbeiter der Anwendungs-Entwicklung in einem einzigen Subnetz zusammenfassen möchte.

Ein wichtiger Aspekt fehlt jedoch noch: Was ist, wenn zwei Clients in unterschiedlichen Access-Netzen direkt miteinander kommunizieren wollen? Wie funktioniert Peer-to-Peer-Kommunikation, wie z.B. bei Voice und Video? Hier würde der Client nicht das Default Gateway ansprechen, sondern stattdessen per ARP (Address Resolution Protocol) oder mit dem IPv6-Pendant Neighbor Discovery Protocol (NDP) direkt nach der IP-Adresse des Kommunikationspartners suchen, um dessen MAC-Adresse herauszufinden. In diesem Fall muss der lokale xTR an Stelle des Client antworten, wollte man nicht ARP Requests oder Neighbor Solicitations in alle anderen Access-Netze übertragen oder mit Hilfe des Map Resolvers auflösen. Der Router wird damit zum ARP bzw. NDP Proxy.

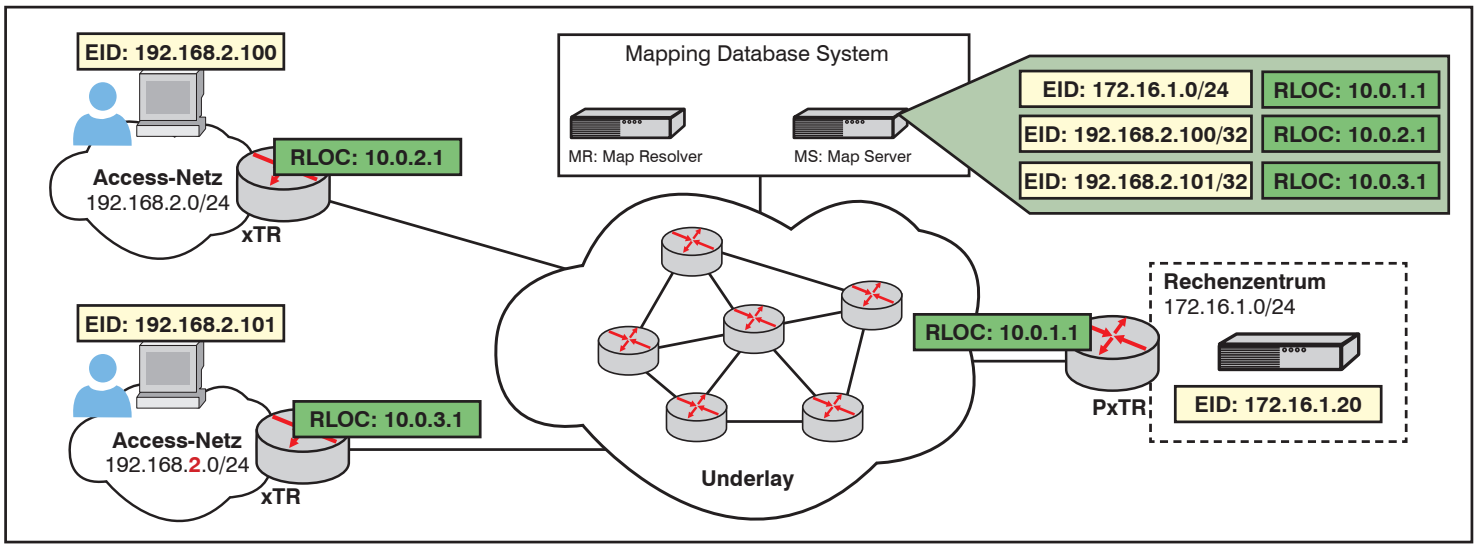

Abbildung 3: Grundzüge der Realisierung von LISP in der Cisco Campus Fabric

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

Das Feature heißt bei Cisco „Local Proxy ARP“, da hierbei auf die IP-Adressen des angeschlossenen Subnetzes geantwortet wird und nicht auf alle anderen (wie normalerweise bei Proxy ARP).

Cisco Campus Fabric emuliert also Layer-2-Netze, was die Wahl des Begriffs „Fabric“ erklärt. Damit bezeichnen Hersteller Gebilde aus Switches, die übergreifend eine Layer-2-Konnektivität bereitstellen, als handele es sich um einen einzigen ausgedehnten Layer-2 Switch. Genau genommen handelt es sich zwar nur um “Private VLANs”, die dank Local Proxy ARP am Ende doch miteinander kommunizieren können, aber das tut letztlich nichts zur Sache.

LISP an sich ist jedoch eine Routing-Technik. Insbesondere lassen sich in LISP nur IP-Pakete enkapsulieren. Der RFC 6830 nennt explizit die beiden Adress-Familien IPv4 und IPv6 und nicht mehr. Insbesondere keine MAC-Pakete. Das mag der Grund dafür sein, dass Cisco sich gegen das Enkapsulieren in LISP entschieden hat und stattdessen VXLAN gewählt hat. Abbildung 4 verdeutlicht den Unterschied. Enkapsuliert man ein Paket in LISP, geht der MAC Header verloren. In VXLAN bleibt er erhalten.

Allerdings wird der ursprünglich MAC Header bereits dadurch verändert, dass der ITR das Paket entgegennimmt. Jeglicher IP-Router überträgt eben nur IP-Pakete, ohne MAC Header. Für die Enkapsulierung hätte Cisco also grundsätzlich auch LISP wählen können, ohne die Funktionsweise der Campus Fabric zu beschränken – jedenfalls soweit ich sie bis hierher dargestellt habe. Vielleicht liegt der Grund für die Wahl der VXLAN-Enkapsulierung darin, dass im VXLAN Header etliche Bits als „reserved“ markiert sind. Sie sind also frei für spätere Nutzung. Und tatsächlich trägt Cisco dort eigene Informationen ein. Vielleicht liegt der Grund auch einfach in der Hardware begründet. Damit eine hohe Performance erzielt wird, sollten Switches Pakete mit ihrer Hardware weiterleiten und entsprechend modifizieren. Und für VXLAN, also die VXLAN Tunnel Endpoints (VTEP), hat Cisco bereits aus anderen Projekten entsprechende Implementierungen, die sich nun ohne weiteres für Campus Fabric nutzen lassen.

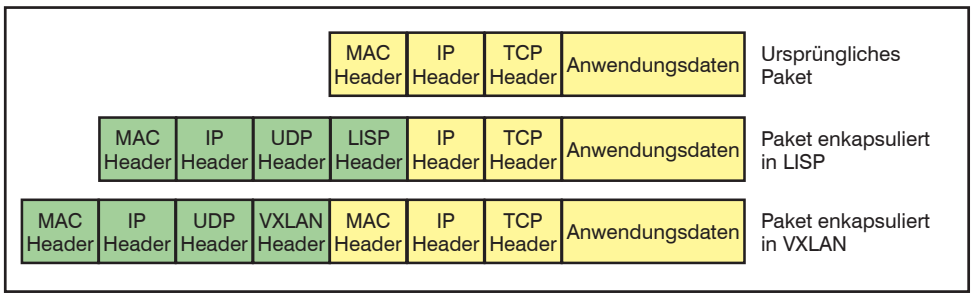

Abbildung 4: Enkapsulierung eines Pakets in LISP und VXLAN

Implementierung von Cisco Campus Fabric

Cisco unterstützt Campus Fabric auf seinen bekannten LAN Switches und Routern. Wesentliche Elemente der Lösung konnten also offensichtlich in Software realisiert werden, so dass in relativ kurzer Zeit eine recht breite Produktunterstützung gegeben ist. Folgende Elemente gibt es in der Campus Fabric (siehe Abbildung 5):

- Edge Nodes sind die Access Switches, mit denen die Access-Netze realisiert werden. Edge Nodes enthalten die LISP Router xTR und sind letztlich auch Teil des gerouteten Underlay. Es handelt sich also um typische Access Switches mit Layer-3-Funktionalität. Darüber hinaus ist die Fähigkeit der Port-basierten Zugangskontrolle (Network Access Con-

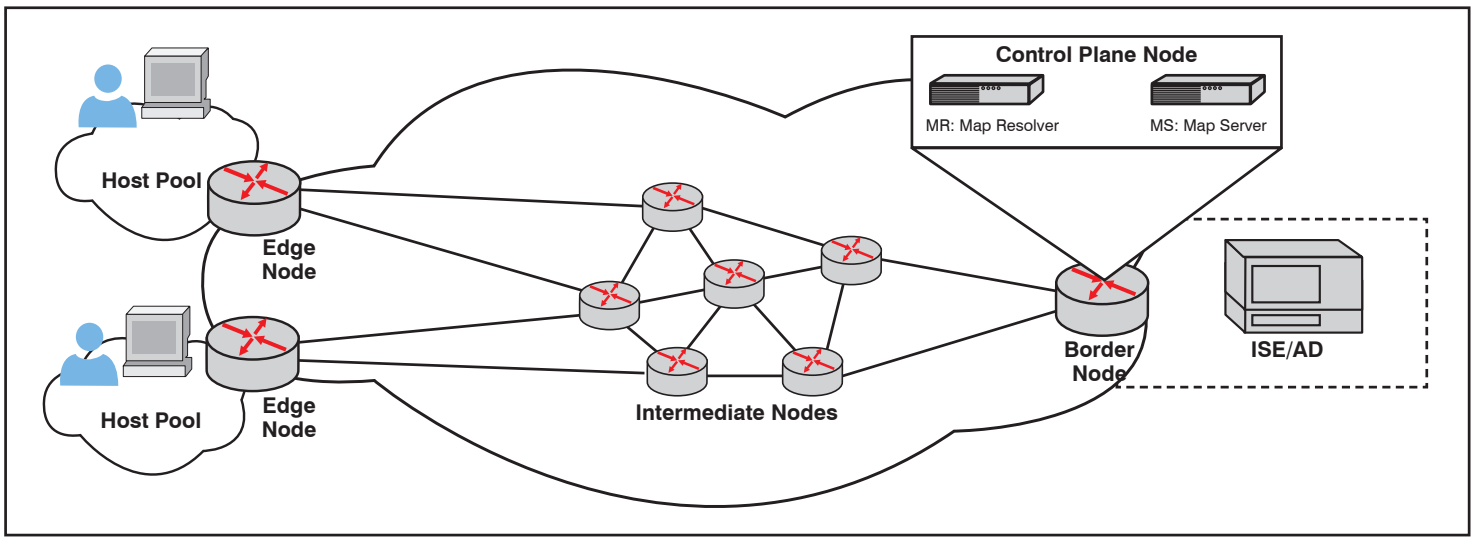

Abbildung 5: Elemente der Cisco Campus Fabric

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

trol, NAC) für Campus Fabric nützlich, dazu später mehr. Als Edge Nodes eignen sich einige modulare und nicht-modulare High-End Access Switches von Cisco.

- Intermediate Nodes sind Standard Layer-3 Distribution und Core Switches ohne zusätzliche Anforderungen.

- Die Kopplung der Campus Fabric in Richtung von Rechenzentrums-Netzen wird von den so genannten Border Nodes realisiert. Als Border Nodes sind die großen modularen Switches aus dem Campus- und RZ-Portfolio geeignet. Aber auch bestimmte Router und nicht-modulare High-End Access Switches können diese Funktion wahrnehmen.

- Map Resolver und Map Server werden als Control Plane bezeichnet. Diese Funktionen sind in der Switch Software realisiert. Die meisten der als Border Nodes geeigneten Switches können auch Control Plane Nodes sein.

- Die von der Control Plane verwaltete Datenstruktur bezeichnet Cisco als Host Tracking Database (HTDB).

Wie bereits erwähnt, wird NAC im Zusammenhang mit Campus Fabric unterstützt. Das wird sogar als der wesentliche Vorteil der Lösung angesehen. Nach Authentisierung am Switch Port wird der Client einem Host Pool zugewiesen. Darunter versteht Cisco die Access-Netze, die zu einer Benutzergruppe gehören. Alle Clients einer Benutzergruppe teilen sich also eine Fabric. Der Border Node routet die Pakete der Fabric in ein entsprechendes virtuelles Netz (VRF), über das letztlich die Server erreicht werden.

Der Charme daran ist aus meiner Sicht die einfache Konfiguration. Alle Edge Nodes können nämlich gleich konfiguriert werden. Wird ein zusätzlicher Edge Node benötigt, kopiert man die Konfiguration von einer Vorlage und passt die Loopback-Adresse für das Routing im Underlay an. Es brauchen keine VRF oder VLANs eingerichtet zu werden. Switch Virtual Interfaces (SVIs) haben auf allen Switches dieselben IP-Adressen (Anycast Gateway).

Sozusagen als Gegenleistung für diese Einfachheit muss man die Security-Lösung „TrustSec“ von Cisco implementieren, die auf der Identity Services Engine (ISE) als zentralem Element basiert. TrustSec nutzt zur Klassifizierung von Daten unter anderem ein proprietäres Informations-Element, das Scalable Group Tag (SGT), das im VXLAN Header mit übertragen wird. Letztlich ermöglicht die Tatsache, dass Pakete durch das Underlay getunnelt werden, also erst die Übertragung des SGT und somit die Qualifizierung des Datenverkehrs an zentralen Firewall-Elementen, ohne dass man wie bisher IP-Adressen und TCP-Portnummern dafür heranziehen müsste. Und damit sich das alles noch verwalten lässt, bietet Cisco mit dem DNA Center [1] gleich die passende Management-Lösung mit an.

Weitere Aspekte

Campus Fabric ist, wie gesagt, ein Overlay. Pakete werden enkapsuliert über ein Underlay übertragen. Durch die Enkapsulierung werden die Pakete länger (vgl. Abbildung 4). Diesen Effekt kennen Sie von allerlei VPN-Techniken. Normalerweise bekommt man davon wenig mit, weil die enkapsulierenden Elemente Router sind. Übliche Betriebssysteme unterstützen zu diesem Zweck das Verfahren der Path MTU Discovery (RFC 1191 bzw. RFC 8201 für IPv6) , welches die maximale Paketlänge (Maximum Transmission Unit) auf dem Weg zum Ziel bestimmt. Hierzu wird das Don’t Fragment Bit im IP Header gesetzt, und der Router teilt dem Absender mit, wenn er das Paket verwerfen musste.

Im Gegensatz dazu emuliert die Campus Fabric eine Layer-2-Technik. Dennoch scheint Path MTU Discovery unterstützt zu werden, schließlich sind alle Edge Nodes auch Router. Cisco empfiehlt jedoch, dass im Underlay eine größere MTU eingestellt wird als die im Ethernet normalerweise verwendeten 1500 Bytes. Es sollten also auf allen Switches des Underlay wie auch auf den Border und Edge Nodes Jumbo Frames aktiviert werden.

Das Emulieren von Layer-2-Netzen impliziert eigentlich auch die Unterstützung von Broadcasts und Multicasts. Damit sieht es bei LISP schlecht aus, denn es gibt keine Funktion, zur Verteilung solcher Pakete (LISP ist eine Layer-3-Technik). Cisco ermöglicht es dennoch, Multicasts für bestimmte Host Pools zu aktivieren. Die Verteilung von Multicasts erfolgt durch den ITR an alle ETR. Damit das Paket hierfür im Underlay nicht wiederholt ausgesandt werden muss, wird es per IP Multicast an die ETR verteilt. Das Underlay muss also über die Fähigkeit des Multicast Routing verfügen.

Eine weitere Einschränkung muss bezüglich DHCP hingenommen werden. Normalerweise gibt es in gerouteten Netzen die Funktion des DHCP Relay Agent, die in Routern implementiert ist, die als Default Gateways für Clients wirken. Empfängt der Relay Agent einen DHCP Request als Broadcast, wandelt er ihn in ein Unicast-Paket um, das an die IP-Adresse des DHCP Servers gesandt wird. Gleichzeitig trägt der Relay Agent die IP-Adresse des Interface, auf dem er den Broadcast empfangen hatte, in ein entsprechendes Feld im DHCP Header ein. An dieser Adresse erkennt der DHCP Server, aus welchem Topf („Scope“) er eine IP-Adresse herausnehmen und dem Client anbieten soll.

Dieses Verfahren funktioniert laut Cisco im Zusammenhang mit Campus Fabric nicht. Als Grund wird angegeben, dass als Absenderadresse für DHCP nur die Loopback-Adresse des Edge Node in Frage kommt, die bekanntlich in keinem Zusammenhang zum IP-Netz des Host Pool steht. Stattdessen muss hier die DHCP Option 82 verwendet werden, in die der Relay Agent eine sogenannte Circuit ID einträgt. Damit wird das VLAN bzw. Subnetz bezeichnet, für das der DHCP Server eine IP-Adresse anbieten soll. Selbstverständlich muss dieses Verfahren von Ihrem DHCP Server unterstützt werden, und Sie müssen es entsprechend einrichten. Zuletzt weise ich noch auf einen interessanten Nebeneffekt des Einsatzes von Cisco Campus Fabric auf WLAN hin. Bekanntlich werden WLAN seit Jahren auf Basis von WLAN Controllern aufgebaut. Die Access Points tunneln alle Pakete zum Controller. Damit realisiert der WLAN Controller sozusagen eine WLAN Fabric, also ein virtuelles Layer-2-Netz über alle Access Points. Mobile Endgeräte können sich von Access Point zu Access Point hangeln, ohne ihre IP-Adressen anpassen zu müs-

Der Netzwerk Insider

Systematische Weiterbildung für Netzwerk- und IT-Professionals

sen. Ich habe mich zur Darstellung eines Bildes aus meinem Insider-Artikel vom September 2013 bedient (Abbildung 6). Auf der linken Seite erkennt man das Prinzip des WLAN Controllers mit zentraler Data Plane. Der gesamte WLAN Traffic läuft über den Controller, was diesen an die Grenze seiner Performance bringen kann.

Zur Abhilfe dieses Problems bieten die Hersteller schon seit langem die Option des Local Bridging [2] an. Hierbei leitet der Access Point den Datenverkehr direkt ins LAN. Der WLAN Controller dient einzig noch dem Management der Access Points. Der Nachteil dieses Verfahrens war bisher der Verzicht auf Mobilität, denn WLAN-Endgeräte mussten beim Wechsel zu einem anderen Access Point im Allgemeinen eine neue IP-Adresse anfordern. Nicht so mit einer Fabric (Abbildung 6 rechte Seite). Hier wird das Layer-2 Overlay durch die Fabric bereitgestellt. Alle Access Points können den mobilen Endgeräten dasselbe IP-Subnetz anbieten. Die volle Performance der über das Underlay vermaschten Fabric steht damit auch dem WLAN zur Verfügung. Leider wird diese Variante von Cisco Campus Fabric offensichtlich nicht unterstützt – wie schade. Über die Gründe kann nur spekuliert werden. Möglicherweise dauert der Vorgang der Client-Registrierung durch den ITR zu lange oder man fürchtet eine Überlastung der Control Plane durch häufiges WLAN Roaming.

Zusammenfassung

Cisco stellt mit dem Produkt Campus Fabric eine Technik vor, mit der sich Access-Netze und die zugehörigen IP-Subnetze auf beliebigen Access-Ports eines Campus bereitstellen lassen. Die Bereitstellung lässt sich statisch einrichten oder dynamisch mit Hilfe der Möglichkeiten von Cisco TrustSec. Zur Implementierung können viele Switches des bereits existierenden Cisco-Portfolios eingesetzt werden. Jedoch benötigen die Switches – insbesondere auch die Access Switches – erweiterte Routing Funktionen. Es steht also nicht zu erwarten, dass Campus Fabric eine Low-Cost-Lösung sein wird.

Die Fähigkeit, IP-Subnetze auf beliebigen Access-Ports eines Campus bereitstellen zu können, ist für Netze, bei denen Endgeräte verschiedener Benutzergruppen bzw. Sicherheitszonen sich ein LAN teilen, hilfreich. Die Konfiguration der Netzkomponenten wird dadurch vereinfacht. So lassen sich bei Umzügen benötigte Anschlüsse in kurzer Zeit über ein zentrales Management bereitstellen, was Cisco veranlasst, diese Technik als Software-defined Access (SDA) zu titulieren.

Redundanzfunktionen des der Campus Fabric unterliegenden Netzes basieren auf den bekannten und bewährten Routing-Algorithmen. Entsprechende Konfiguration vorausgesetzt, lassen sich im Fehlerfall schnelle Wiederherstellungszeiten garantieren. Probleme, die in ausgedehnten Layer-2-Netzen durch Schleifenbildung entstehen können, werden durch das Routing-basierte Tunnelprinzip vermieden.

Basis der Datenübertragung (Data Plane) ist Virtual Extensible LAN (VXLAN). Als Control Plane der Campus Fabric hat Cisco das ansonsten recht unbekannte Label/ID Separator Protocol (LISP) gewählt. Genau genommen ist das meines Erachtens die erste wirklich sinnvolle Anwendung von LISP überhaupt. Dennoch stellt sich die Frage, wieso Cisco ausgerechnet dieses Protokoll verwendet. Im Data Center wird mit Ethernet VPN (EVPN) eine Technik mit vergleichbaren Zielen angeboten, die einerseits ebenfalls auf VXLAN basiert, andererseits das Border Gateway Protocol (BGP) als Data Plane verwendet. Außerdem verfügt Cisco mit Fabric Path über eine weitere Layer-2 Fabric. So läge es nahe, eine dieser Techniken auch im Access anzubieten. Cisco wendet ein, dass im Access-Bereich das LISP geeigneter sei, weil hierfür letztlich weniger Rechenleistung auf den Komponenten benötigt werde.

Wie dem auch sei, ich finde die Lösung technisch interessant. Ob allerdings die damit einhergehende enge Bindung an den Hersteller Cisco Systems und seine umfassenden Management- und Security-Lösungen am Ende gerechtfertigt ist, muss von Fall zu Fall abgewogen werden. Ich bin jedenfalls gespannt, ob der Campus Fabric eine breite Akzeptanz gegönnt sein wird.

Erläuterungen

[1] Das DNA Center ist der Nachfolger des Application Policy Infrastructure Controller Enterprise Module (APIC-EM). „DNA“ steht hier für Digital Network Architecture.

[2] Bei Cisco als Flex Connect bezeichnet

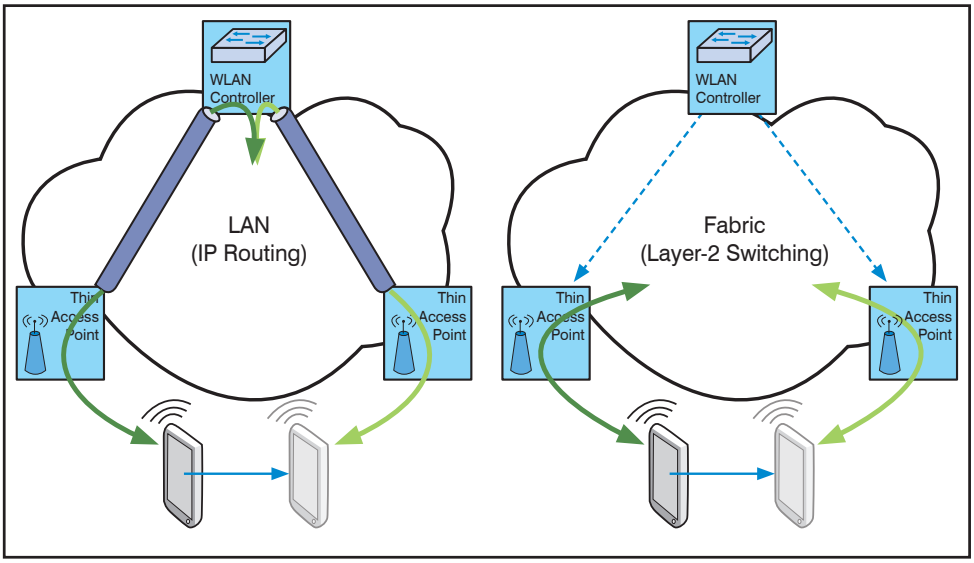

Abbildung 6: WLAN Roaming in Layer-2 Fabrics