BGP-Instabilität identifizieren

Das Netz unter der Lupe

Es gibt viele Elemente, die eine Auswirkung auf die Verfügbarkeit von netzwerkübergreifenden Diensten haben. Daher ist Redundanz in der Topologie – verbunden mit Ausfallsicherheit in der Konfiguration – von entscheidender Bedeutung. Routing-Protokolle dienen dazu, diese Redundanz und das Failover auf Sicherungsverkehrspfaden zu verwalten, falls im aktiven Pfad ein Fehler auftritt. Damit dies erfolgreich funktioniert, ist es wichtig, solche Änderungen in Routing-Konfigurationen im Blick zu behalten, die sich auf die Netzwerkstabilität auswirken.

Das Border Gateway Protocol (BGP – und insbesondere die aktuelle Version 4) kommt zum Einsatz, um Netzwerke miteinander zu verbinden. Damit zum einen sichergestellt ist, dass Subnetze erreichbar sind, und zum anderen redundante und ausfallsichere Pfade zwischen Netzwerken existieren, um die Verfügbarkeit von Diensten auch im Fall eines Fehlers zu gewährleisten. Im BGP wird eine Peering-Beziehung zwischen Routern in separaten autonomen Systemen (AS) hergestellt und Routing-Präfixinformationen ausgetauscht. Diese Daten dienen dann dazu, den „besten“ Pfad zu berechnen. Maßgeblich ist dabei die Qualität des Pfads. Sie lässt sich standardmäßig über die Anzahl der AS kalkulieren, die die Daten für die Übertragung von der Quelle zum Ziel durchqueren müssen.

Wie vermutlich jeder Netzwerkadministrator bestätigen kann, verursacht die Instabilität des BGP-Peerings bei großen Netzwerken alle Arten von Leistungsproblemen. Ereignisse – etwa dauerhafte Verbindungsfehler – können Update-Sequenzen entlang von Pfaden im Netzwerk auslösen, die folgende Auswirkungen haben können: Temporäre Forwarding-Loops für betroffene Netzwerkpräfixe während der Konvergenz, extremen Bandbreitenverbrauch aufgrund erhöhter Kommunikationsaktivität zwischen Netzwerkknoten sowie eine steigende CPU-Auslastung in Netzwerkknoten aufgrund der Notwendigkeit, große Mengen sich ständig ändernder Daten verarbeiten zu müssen. Ein Beispiel: Internet-Routing-Tabellen haben Größen erreicht, sodass sie nie vollständig konvergieren. Dies ist symptomatisch für die kontinuierlichen Aktualisierungen und Löschungen von Routing-Einträgen, die durch Verbindungsereignisse innerhalb und zwischen den AS entstehen, wie sie rund um die Uhr und auf der ganzen Welt auftreten.

Daten von regionalen Internetregistraren – den Organisationen, die für die Zuweisung und Verwaltung von Internetadressen und AS-Nummern verantwortlich sind – zeigen, dass mit zunehmender Größe der Internet-Routing-Tabelle die Anzahl der täglichen Updates schrittweise zunimmt. Dies kann sich auf bis zu einem Viertel der gesamten Einträge in der Routing-Tabelle auswirken.

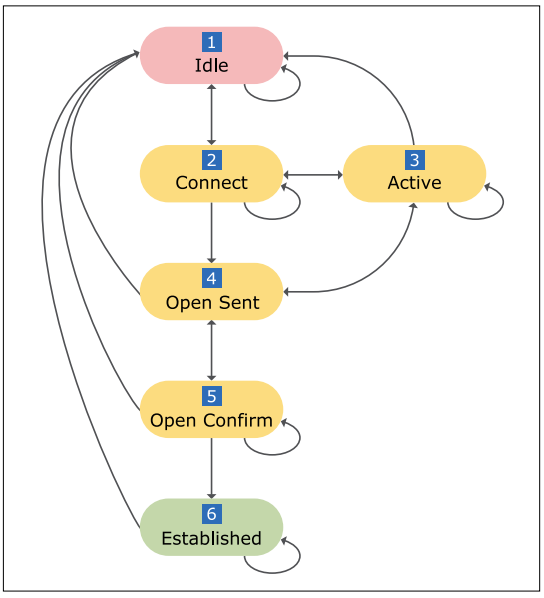

Bild 1. Verschiedene Zustände einer BGP-Verbindung.
Bild: Netcor

Wie stabil ist „stabil“?

BGP dient üblicherweise in einem Unternehmen dazu, separate Netzwerke miteinander zu verbinden. Die Verwaltung geschieht häufig von verschiedenen Organisationen oder verschiedenen Teilen derselben Organisation aus. Daraus folgt, dass die Verbindungen, sobald sie einmal zustande gekommen sind, idealerweise für lange Zeiträume bestehen bleiben sollten. Gibt es dabei Änderungen, kann dies weitreichende Konsequenzen haben. Es lohnt sich daher, die Stabilität dieses Peerings kontinuierlich im Auge zu behalten.

Aber wie genau können Betreiber diese Stabilität messen? Wie können sie insbesondere feststellen, ob und wann das BGP-Peering seinen stabilen Zustand nicht aufrechterhält? Dazu kann man sich auf zwei ganz spezielle Eigenschaften der Peer-Beziehung konzentrieren.

BGP funktioniert, indem es zunächst eine Reihe von Peering-Beziehungen zwischen Router-Paaren herstellt und dann Präfixinformationen über diese hergestellten Verbindungen austauscht. Der Peering-Prozess durchläuft mehrere Phasen in einer bestimmten Reihenfolge, um sicherzustellen, dass beide Nachbarn die zwischen ihnen übertragenen Daten korrekt austauschen und verstehen können (Bild 1). Um die

Stabilität zu bestimmen, lassen sich für jede BGP-Nachbarbeziehung im Netzwerk spezifische Fragen beantworten: Wie ist der aktuelle Stand der Peer-Verhandlungen? Ist die Beziehung vollständig aufgebaut oder verhandelt sie noch in Richtung des Zustands „Established“? Wenn sie „etabliert“ ist, wie lange besteht dann die Peer-Beziehung in diesem Zustand? Ist der Zeitraum kurz, kann dies ein Indikator dafür sein, dass ein regelmäßiger Zustandswechsel stattfindet.

Um die allgemeine Stabilität oder die des BGP-Peerings im Netzwerk manuell zu überprüfen, kann ein Netzwerkadministrator folgende Schritte durchführen:

1. Anmeldung bei jedem Gerät mit Routing-Funktionen (also Router, Layer-3-Switch und Firewall) im Netzwerk,
2. mit Hilfe der herstellerspezifischen CLI-Befehle feststellen, ob BGP-Sitzungen konfiguriert sind,
3. Überprüfung des Status jeder BGP-Peering-Beziehung, die den relevanten Knoten enthält,
4. Festhalten der Ergebnisse in einer Tabelle und
5. (möglicherweise) Weitergabe der Details an einen erfahreneren Netzwerkexperten zur Analyse.

Bild 2. Manuelle Abfrage der BGP-Peers. Bild: Netcor

```
[admin15@L64R7] > /routing bgp peer print
Flags: X – disabled, E – established
 #   INSTANCE        REMOTE-ADDRESS    REMOTE-AS
 0 E default         10.64.109.108     64
 1 E default         10.64.255.101     64
 2   default         10.64.255.44      196652
```

Abhängig von der Größe des jeweiligen Netzwerkes können zwischen zehn und mehreren Hundert von BGP-Peering-Beziehungen bestehen, deren Überprüfung somit eine sehr arbeitsintensive Aufgabe darstellt. Außerdem ist ein erfahrener Netzwerkexperte nötig, um feststellen zu können, wo Instabilität auftreten kann und wo die Ressourcenverfügbarkeit möglicherweise eingeschränkt ist.

Automatisierung der Administrationsprozesse

Netzwerkadministratoren erlernen von den meisten Geräteherstellern diverse Automatisierungstechniken für die Überprüfung der Zustände von BGP-Sitzungen. Die Hersteller stellen zum Beispiel Management-Lösungen zur Verfügung, die auf den Netzwerkgeräten nach Herstellervorgaben eine Template-basierende Konfiguration anwenden. Im Folgenden wird der Status der Geräte über Dashboards und Widgets überwacht, um sicherzustellen, dass alles ordnungsgemäß funktioniert. Bei diesen Konfigurationen handelt es sich gewöhnlich um herstellerspezifische Plattformen, die zu einer bestimmte Netzwerklösung passen (etwa Cisco ACI für das Datenzentrum oder Velocloud für das softwaredefinierte WAN). Neuere Betriebssysteme für Netzwerkgeräte sind so aufgebaut, dass sie sich sowohl über APIs als auch über eine herkömmliche Befehlszeilenschnittstelle (CLI) verwalten lassen. Daher ist es möglich, mit Hilfe von Programmiertechniken, Konfigurationen zu verteilen und zu überwachen.

In diesem Fall kann der Netzwerkexperte Skripte schreiben, um die gewünschten Aktionen auszuführen und den aktuellen Status zu überprüfen. Er muss dafür jedoch das Programmieren lernen, die Skriptentwicklung verwalten und aufrechterhalten sowie für jeden Hersteller und jede Gerätefamilie die Skripte anpassen.

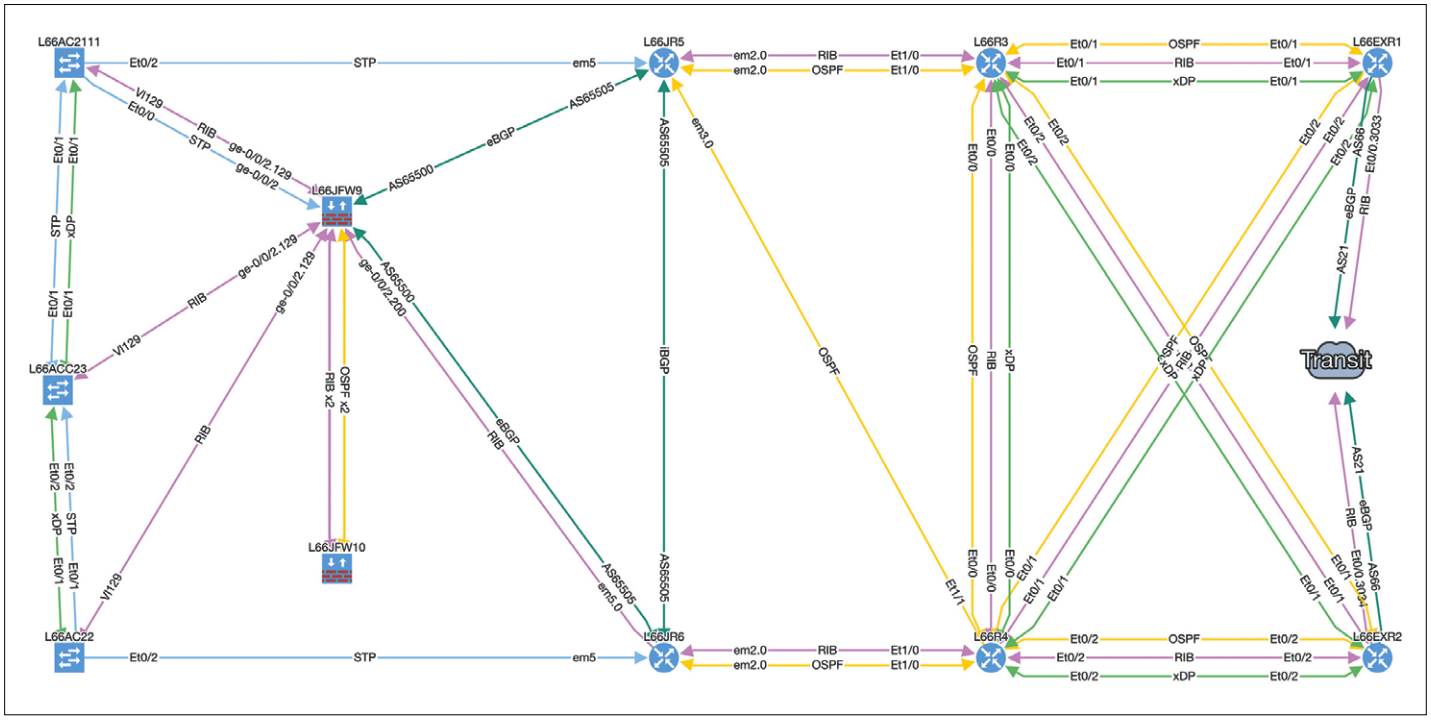

Bild 3. Automatisch erstelltes Netzwerkdiagramm der aktiven Routing-Verbindungen. Bild: Netcor

Name	Column	Results			
BGP Session Age	Uptime	125	28	12	28
BGP Neighbor State	State	163	0	16	14
BGP Received Prefixes	#Rcvd prefixes	133	28	32	

BGP Neighbors

Hostname	Site	Local AS	Local VRF	Local IP	Neighbor	Neighbor ID	Neighbor IP	Neighbor AS	Type	State	#Rcvd prefixes	Neighbor Interface	Uptime
L64R7	L64	64	main	(empty)	L64R4	0.0.0.0	10.64.255.44	3.44	external	active	0	bridge-AS65539.4344	0 seconds
L71FW9/ipf	L71	65000.71	(empty)	0.0.0.0	L71EXR1	0.0.0.0	10.71.111.101	71	external	active	0	Et0/3	0 seconds
L1LB1	L1	64580	(empty)	0.0.0.0	L1XOS1	0.0.0.0	10.241.1.108	65000.747	external	idle	0	VLAN100	0 seconds
L1LB1	L1	64580	(empty)	0.0.0.0	L1R16	0.0.0.0	10.241.255.16	64580	internal	connect	0	lo0.0	0 seconds
L77R12-LEAF6	L77	1	(empty)	(empty)	L77RS-SPINE1	10.77.255.5	10.77.255.5	1	internal	idle	7	Lo0	0 seconds
L71FW9/root	L71	65000.71	(empty)	0.0.0.0	L71R3	0.0.0.0	10.71.110.103	71	external	active	0	Et1/3	0 seconds
L77R10-LEAF4	L77	1	VRF1	(empty)	L77RS-SPINE1	0.0.0.0	10.77.255.5	1	internal	active	0	Lo0	0 seconds
L1R6.lab.ipfabric.io	L1	0	default	0.0.0.0	L1R3	0.0.0.0	10.241.255.3	2	external	idle	0	Lo0	0 seconds
L77R11-LEAF5	L77	1	(empty)	(empty)	L77RS-SPINE1	10.77.255.5	10.77.255.5	1	internal	idle	7	Lo0	0 seconds
L1R11	L1	64560	VRFXX	(empty)	L1R10	0.0.0.0	172.16.16.1	64544	external	idle	0	Fa2/0	0 seconds
L77R9-LEAF3	L77	1	VRF1	(empty)	L77RS-SPINE1	0.0.0.0	10.77.255.5	1	internal	active	0	Lo0	0 seconds
L45R4	Ballpark	64545	(empty)	(empty)	L45XR11	0.0.0.0	10.45.255.111	64545	internal	idle	0	Lo0	0 seconds
L1fw8	L1	64580	(empty)	(empty)	L1R16	10.241.255.16	10.241.255.16	64580	internal	established	2	lo0.0	0 seconds
L1fw8	L1	64580	(empty)	(empty)	L1XOS1	10.241.1.108	10.241.1.108	65000.747	external	established	14	VLAN100	0 seconds
L1Quagga1	L1	65534	0	(empty)	L1R4	0.0.0.0	10.241.255.4	64580	external	active	0	Lo0	0 seconds
L1R4	L1	64580	(empty)	10.241.255.4	L1FW1	0.0.0.0	10.241.1.104	64581	external	connect	0	Gi0/0	0 seconds

Bild 4. Statusbericht über die Zustände der BGP-Sessions. Bild: Netcor

Für die Identifizierung der Ursache von BGP-Instabilität lässt sich dieser manuelle Prozess auch automatisieren. Dies kann viel Zeit sparen, indem ein automatisierter Prozess zum Abrufen der Daten erstellt wird, die einem Netzwerkexperten bei der Behebung des Problems helfen. In das Skript lässt sich sogar eine einfache Auswertung der Daten einbauen, um die Situation noch genauer zu analysieren. Dies setzt jedoch voraus, dass das Problem bereits bekannt ist und dass das Skript auch abläuft, um die Informationen für die Behebung des entsprechenden Problems zu sammeln.

Ein alternativer Ansatz

Eine effizientere Methode ist die Verwendung einer automatisierten Netzwerkinfrastruktur-Management-Plattform, die das Netzwerk intelligent untersucht, indem sie regelmäßig die gesamte Konfiguration und den Betriebszustand aller Netzwerkknoten erfasst und analysiert. Sobald diese Daten an einem zentralen Ort in anbieterneutraler Form gespeichert sind, sind sie auf eine Reihe von Bedingungen und Problemen hin auswertbar. Dann lassen sie sich proaktiv auf einem Dashboard zusammenfassen oder mittels Alarmes an ein übergeordnetes Management-System weiterleiten.

Mit einem solchen System kann der Betreiber eine beliebige Anzahl von Konfigurations- oder Betriebselementen auf ihre korrekte Funktionalität hin einschließlich der oben beschriebenen BGP-Stabilität überprüfen. Durch den Überprüfungsprozess ist es dann auch möglich, zuvor verborgene Inkonsistenzen und Schwachstellen im Netzwerk aufzudecken. Die Option, diese Analyse in Umgebungen mit mehreren Komponenten unterschiedlicher Hersteller durchzuführen, trägt dazu bei, eine einheitliche Ende-zu-Ende-Ansicht des Netzwerks zu erhalten.

Wichtig ist dabei die Möglichkeit, diese Daten sowohl in Tabellen als auch in einer Topologiekarte über eine intelligente, zentrale Web-Benutzeroberfläche anzuzeigen. Ebenso wichtig ist es, dass diese Daten dann über eine API anderen Management-Plattformen zugänglich sind.

IP Fabric ist zum Beispiel ein solches System. Es stehen sofort zahlreiche Validierungsprüfungen zur Verfügung, mit denen ein Techniker von der Konfigurationskonsistenz bis hin zur Topologieredundanz alles überprüfen kann. Die Stabilität der Routing-Protokolle spielt bei diesen Überprüfungen eine wichtige Rolle. Dazu gehören auch die Tests, die zuvor in Bezug auf das BGP genannt sind und deren Ergebnisse sich in einem einfachen Dashboard im Ampelstil hervorheben lassen. Die intelligenten Netzwerkdiagramme können dann verwendet werden, um sehr schnell einen Einblick in die Beziehungen zwischen Knoten im Netzwerk von Layer 1 bis Layer 3 und Routing-Protokollen zu erhalten und den Problemen schnell und effizient auf den Grund zu gehen.

Nur mit einem spezialisierten Management-System ist eine automatisierte, herstellerübergreifende Überwachung der Netzwerkinfrastruktur auf ihre korrekte Funktionalität hin möglich. Mit dem intelligenten Netzwerkerkennungsprozess von IP Fabric ist der Netzwerkadministrator in der Lage, große Netzwerke innerhalb von Minuten zu durchleuchten.

Diese Ergebnisse liefern Auskunft über Abweichungen von definierten Normalzuständen und unterstützen den Netzwerksupport bei der schnellen Identifizierung von Inkonsistenzen und Problemen, wenn beispielsweise BGP-Peering die Netzstabilität beeinflusst. Außerdem steht eine ständig aktualisierende Netzwerkdokumentation zur Verfügung. Im Fall von Leistungsproblemen kann das System auch den Ende-zu-Ende-Pfad zwischen den beteiligten Systemen in einem Netzwerkdiagramm darstellen.

Jos Op 't Root/jos

Jos Op 't Root ist Geschäftsführer von Netcor, www.netcor.de.