

Temporäre Schleifen während der Konvergenz von IGP-Routing-Protokollen

Johannes Formann

30. Juli 2007

In den aktuellen IGP-Protokollen kann es während der Konvergenzzeiten zu ungewollten Routingschleifen kommen, die die Funktionsfähigkeit des Netzes beeinträchtigen.

An einem Beispiele wird vorgestellt, wie es dazu kommen kann, und Ansätze vorgestellt, die diese Schleifen und Auswirkungen vermeiden oder zumindest reduzieren.

Inhaltsverzeichnis

1	Einleitung	1
2	Kurzvorstellung Link-State-Protokolle	1
3	Entstehung der Konvergenzzeit	1
4	Beispiel für die Entstehung einer Schleife bei OSPF	2
5	Ansätze zur Vermeidung/Behandlung von Routingschleifen	3
5.1	Paketweiterleitung bei geplanten Neustart eines Routers	3
5.2	Geordnete Aktualisierung	4
5.3	Schnittstellenspezifische Weiterleitungstabellen	4
6	Zusammenfassung der Vor- & Nachteile	4

Abbildungsverzeichnis

1	Topologie für das Beispiel	3
---	--------------------------------------	---

Tabellenverzeichnis

1	Routing Ereignisse bei OSPF nach Ausfall des Links E-D	3
2	Routing Ereignisse bei den verschiedenen Verfahren	5

1 Einleitung

Wenn durch eine inkonsistente Sicht der aktuellen Netztopologie¹ mehrere Router IP-Pakete im Kreis weiterleiten, dann spricht man von einer Routing-Schleife.

Routing-Schleifen können ein großes Problem in Netzwerken darstellen, dadurch können auch breitbandige Verbindungen überlastet werden. Dadurch werden weitere Teile des Netzes beeinträchtigt.

Während die Probleme mit den Routing-Schleifen während der Konvergenzzeit bei Distance-Vector-Routingprotokollen in der Literatur bekannt sind[5, 4, 6], findet sich in der Literatur für Link-State-Routingprotokolle folgende in der Praxis falsche Aussagen:

„*Bessere Konvergenz*: [...] Alle Router innerhalb der Domain aktualisieren augenblicklich ihre Routing-Tabellen.”[4]

„Schnelle Konvergenz ohne Schleifen”[5]

Im folgenden werde ich zuerst die Funktionsweise von Link-State-Protokollen vorstellen, um dann zu erklären woher in der Praxis, die theoretisch nicht vorhandene, Konvergenzzeit kommt. Danach werde ich exemplarisch an OSPF zeigen wie in dem Zeitraum eine Routing-Schleife entstehen kann und drei vorgeschlagene Lösungsansätze präsentieren.

2 Kurzvorstellung Link-State-Protokolle

Bei den Link-State-Routingprotokollen besitzt jeder Router eine vollständige Datenbank der Netztopologie. In dieser Datenbank stehen sämtliche Verbindungen, welche die Router der OSPF Domain haben, mitsamt der zugeordneten Metrik.

Die Metrik beschreibt wie „gut“ eine Verbindung ist, sie kann z.B. die Bandbreite einer Verbindung angeben, oder eine vom Administrator gesetzte Konstante. Sie muss in der Domain die gleiche Bedeutung haben, damit jeder Router mit Hilfe dieser Angabe den besten Pfad berechnen kann.

Wenn ein Router eine Veränderung in der Netzwerktopologie² feststellt, dann wird eine Aktualisierungsnachricht an alle angeschlossenen Router gesendet, die diese auch weiterverbreiten³, so dass alle Router der Domain die Änderung mitbekommen, und ihre Datenbank & Weiterleitungstabellen aktualisieren können.

Weiterführende Informationen finden sich in der Literatur[5].

3 Entstehung der Konvergenzzeit

Die Konvergenzzeit besteht aus mehreren Komponenten:

¹Das heißt, die Router haben unterschiedliche „Meinungen“ darüber, welcher Router das Paket am besten weiterleiten kann

²Eine Verbindung ist ausgefallen/hinzugekommen. Die Metrik einer Verbindung hat sich geändert.

³Das so-genannte Flooding-Protokoll

Fehlererkennungszeit: Die Zeitspanne die notwendig ist, damit benachbarte Router erfahren, dass der Link/Router ausgefallen ist. Muss bei verschiedenen Routern nicht identisch sein.

Während dieses Zeitraumes kann noch keiner der Router den Ausfall erkennen, und die Routingtabellen ändern, so dass Pakete, die über diesen Weg laufen, verloren gehen.

Meldungsweiterleitungszeit: Die Zeitspanne die vergeht, um die Meldung über die Topologieänderung an den nächsten Router weiterzugeben.

Besteht aus dem Zeitraum, der zur Erstellung der Nachricht benötigt wird, zuzüglich evtl. im Protokoll vorgesehenen Wartezeiten um mehrere Meldungen zusammenfassen zu können und der Übertragungszeit.

Diese Zeitspanne akkumuliert sich, je mehr Stationen der Router von dem Router der den Ausfall gemeldet hat entfernt ist.

Routing-Berechnungs & Aktualisierungszeit: Die Zeitspanne die zum Berechnen und Aktualisieren der Weiterleitungstabellen vergeht. Beinhaltet auch evtl. vorgesehene Wartezeiten, die dazu dienen sollen, dass mehrere dicht aufeinander folgende Meldungen zusammen verarbeitet werden können.

Die Dauer hängt von der Implementierung des Protokolls, der verwendeten Hardware, der Anzahl der Links und Router im Netzwerk ab, und kann damit stark variieren.

Diese Zeiten zusammengenommen, können durchaus mehrere Sekunden in einem größeren Netz ergeben[2], so dass die Auswirkungen von evtl. entstehenden Schleifen sehr störend sein können.

Im weiteren Verlauf werde ich für die Beispiele, soweit nicht anders angegeben, folgende Zeiten für alle Geräte annehmen:

Fehlererkennungszeit: 50ms

Meldungsweiterleitungszeit: 100ms

Routing-Berechnungs & Aktualisierungszeit: 400ms für Router die vom Ausfall betroffen sind, und 100ms für Router die nicht vom Ausfall betroffen sind.

4 Beispiel für die Entstehung einer Schleife bei OSPF

In der Abbildung 1 ist die Topologie dargestellt, die für das folgende Beispiel verwendet wird. Es handelt sich dabei um ein Netz welches OSPF nutzt. Die Zahlen an den Kanten geben die Metrik der Verbindung an.

In dem Beispiel (nach [7]) wird gezeigt, wie sich die Pakete von Router-A nach Router-D weitergeleitet werden, wenn die Verbindung zwischen Router-E und Router-D ausfällt.

In der Tabelle 1 sind die Abläufe zu sehen, die in diesem recht kleinen Beispielnetz zu einer Routing-Schleife von 0,1 Sekunden führen. In größeren Netzen ist es möglich, dass die Schleifen länger dauern oder an mehreren Stellen auftreten.

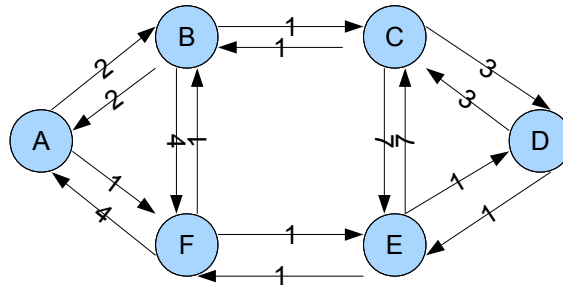


Abbildung 1: Topologie für das Beispiel

Zeit	Ereignis	Pakete von A nach D
0,00s	Ausfall Link E-D	A-F-E-verwerfen
0,05s	D,E: erkennen Ausfall	A-F-E-verwerfen
0,15s	C,F: über Ausfall informiert	A-F-E-verwerfen
0,25s	A,B: über Ausfall informiert	A-F-E-verwerfen
0,35s	B: aktualisiert Routen	A-F-E-verwerfen
0,45s	D,E: aktualisieren Routen	A-F-E-F...(Schleife)
0,55s	C,F: aktualisieren Routen	A-F-B-C-D
0,65s	A: aktualisiert Routen	A-B-C-D

Tabelle 1: Routing Ereignisse bei OSPF nach Ausfall des Links E-D

5 Ansätze zur Vermeidung/Behandlung von Routingschleifen

5.1 Paketweiterleitung bei geplanten Neustart eines Routers

Im folgenden geht es um ein Verfahren, das nicht direkt Schleifen im Netz vermeidet. Es stellt aber einen Ansatz dar, mit dem die Routing-Updates, die durch die Wartung der Router entstehen⁴, vermieden werden. Da diese einen Großteil der Änderung in einem Netz verursachen können[2], wird durch die geringere Häufigkeit von Konvergenzen indirekt die Gefahr von Schleifen geringer.

Dazu wird in [1] vorgeschlagen, die bei vielen modernen Routern vorhandene Trennung zwischen der Paketweiterleitung (oft in Hardware) und der Routenberechnungssoftware zu nutzen. Bei einem geplanten Neustart bleibt der Router (teilweise) funktionsfähig, indem die Paketweiterleitung aufrechterhalten wird, während die Software neu startet.

Damit nun keine Schleifen oder „schwarze Löcher“ entstehen, die dadurch verursacht werden, dass der Router keine Updates an seiner Routing Tabelle ausführt, wenn sich die Topologie woanders im Netz ändert, und damit auch nicht die HELLO-Timer⁵ ablaufen, „meldet er sich ab“ bevor die Software heruntergefahren wird. Die anderen Router im

⁴z.B. zum Einspielen von Updates

⁵für periodische Nachrichten zwischen den Routern, die u.A. dazu dienen festzustellen ob der andere Router noch läuft

Netz können bei der Routenberechnung berücksichtigen, dass die Routen bei diesen Router statisch bleiben, und so „schwarze Löcher“ und Schleifen als Folge davon vermieden werden können.

5.2 Geordnete Aktualisierung

Das Verfahren der geordneten Aktualisierung[3] stellt einen Ansatz dar, bei dem bei einem Update der Weiterleitungstabellen Schleifenfreiheit garantiert werden kann.

Dazu wartet jeder Router der eine Nachricht über eine Änderung bekommt, bis seine Nachbarn, denen er diese Nachricht weitergeleitet hat, bestätigen, dass Sie ihre Weiterleitungstabellen aktualisiert haben. Dabei wird das Netz von „Außen“ nach „Innen“ (wo die Änderungsursache liegt) aktualisiert,.

Durch das Warten auf die „außen“ liegenden Router erhöht sich aber die Konvergenzzeit, zum Teil signifikant[7].

5.3 Schnittstellenspezifische Weiterleitungstabellen

Bei dem Verfahren der schnittstellenspezifischen Weiterleitungstabellen[7] werden die, bei den meisten Routern heutzutage aus Performancegründen schon vorhandenen, schnittstellenspezifischen Weiterleitungstabellen modifiziert, um Pakete die eine Schleife bilden würden zu verwerfen.

Dazu werden mit einen ausgewählten Algorithmus⁶ die Interfaces bestimmt, auf denen Pakete zu Zielen, bei der aktuell im Router gespeicherten Netz-Topologie, nicht ankommen. Bei diesen Interfaces wird dann die Routing-Tabelle so modifiziert, dass diese Pakete verworfen werden⁷.

Da bei mindestens einem der an einer Schleife beteiligten Router die IP-Pakete an einem untypischen Interface ankommen müssen, wird die Schleife so verhindert.

6 Zusammenfassung der Vor- & Nachteile

Es gibt eine Reihe von mögliche Kriterien, nach denen die drei vorgestellten Verfahren verglichen und bewertet werden können. Im Folgenden werde ich für die drei Verfahren erst die Vor- & Nachteile kurz vorstellen, und dann Abwägen für welche Szenarien welches Verfahren geeignet ist.

Für das Verfahren der Paketweiterleitung bei geplanten Neustart eines Routers spricht vor allen, dass in vielen Fällen eine unnötige Routenänderung vermieden wird, und so die Anzahl der Fällen, in denen das Netz konvergieren muss, deutlich reduziert. Dagegen spricht, dass es nur bei einigen bestimmten Hardware-Architekturen verwendet werden kann, eine Änderung des Routing-Protokolls erfordert und im Fall eines ungeplanten Ausfalls keinen Gewinn bringt.

⁶PIPO: Ein- & Ausgehendes Interface wären Identisch

CYCL: Der vorhergehende Router ist in dem Routingpfad enthalten.

NOFP: Der vorhergehende Router ist gleich „weit“ vom Ziel entfernt

⁷Je nach Unterstützung der Hardware explizit oder durch eine Nullroute

		OSPF	Interface spez.	geordnet
Zeit	Ereignis	Pakete: A nach D	Pakete: A nach D	Ereignisse
0,00s	Ausfall Link E-D	A-F-E-verwerfen	A-F-E-verwerfen	Ausfall Link E-D
0,05s	D,E: erkennen Ausfall	A-F-E-verwerfen	A-F-E-verwerfen	D,E: erkennen Ausfall
0,15s	C,F: über Ausfall informiert	A-F-E-verwerfen	A-F-E-verwerfen	C,F: über Ausfall informiert
0,25s	A,B: über Ausfall informiert	A-F-E-verwerfen	A-F-E-verwerfen	A,B: über Ausfall informiert
0,35s	B: aktualisiert Routen	A-F-E-verwerfen	A-F-E-verwerfen	B: aktualisiert Routen
0,45s	D,E: aktualisieren Routen	A-F-E-F...(Schleife)	A-F-E-F-verwerfen	
0,55s	C,F: aktualisieren Routen	A-F-B-C-D	A-F-B-C-D	C: aktualisiert Routen
0,65s	A: aktualisiert Routen	A-B-C-D	A-B-C-D	A: aktualisiert Routen
1,05s				D: aktualisiert Routen
1,15s				F: aktualisiert Routen
1,65s				E: aktualisiert Routen

Tabelle 2: Routing Ereignisse bei den verschiedenen Verfahren

Für das Verfahren der geordneten Aktualisierung spricht, dass es sicher die Bildung von Schleifen verhindert, aber es weist wie in Tabelle 2 sichtbar, deutlich längere Konvergenzzeiten auf, und das Routing-Protokoll muss angepasst/erweitert werden.

Für das Verfahren der schnittstellenspezifischen Weiterleitungstabellen spricht, dass es keine Änderung des Routingprotokolls benötigt und dass es je nach Konfiguration die meisten/alle Routing-Schleifen neutralisiert. Es ist aber nicht auf allen Architekturen einsetzbar, da es auf eine Weiterleitungstabelle pro Schnittstelle angewiesen ist, und auch bei nicht an einer Schleife beteiligten Verbindungen während der Konvergenzphase zu Paketverlusten führen kann.

Das Verfahren der Paketweiterleitung bei geplanten Neustart eines Routers erscheint mir eine gute Ergänzung zu den anderen Verfahren, für den Fall, dass in dem Netzwerk die meiste Router-Hardware das Verfahren unterstützen, und es für *alle* Router entsprechende Software gibt, die diese Erweiterung implementieren.

Da in aktuellen IP-Netzen die Konvergenzzeiten eine große Rolle spielen[2], und auch schon Vorschläge gemacht wurden wie diese auf den (Sub-)Sekundenbereich zu bringen[1], scheint mir der Ansatz der geordneten Aktualisierungen mit den höheren Konvergenzzeiten für die Mehrheit der Netze nicht relevant. Der Ansatz der geordneten Aktualisierungen ist aber für Fälle interessant, bei denen die längere Konvergenzzeit toleriert werden kann, Schleifen vermieden werden müssen, und auch das Verwerfen von Paketen die sonst weitergeleitet werden würden.

Die interessanteste Variante ist, in meinen Augen, die der schnittstellenspezifischen Weiterleitungstabellen, da dabei nur die Software angepasst werden muss. Das Protokoll und die Konvergenzzeit bleibt dabei konstant. Dadurch kann man auch eine langsame Migration zu diesem Verfahren machen, und je höher die Anzahl der Router ist, die dieses Verfahren implementieren, desto geringer wird die Wahrscheinlichkeit von Schleifen. Diese Einfachheit erkaufte man sich aber mit einer leicht erhöhten Paketverlustrate während der Konvergenz.

Literatur

- [1] C. Alaettinoglu, V. Jacobson og H. Yu. Towards milli-second igp convergence. Internet Draft, 2000.
- [2] N. Dubois, B. Fondeviole og N. Michel. Fast convergence project. RIPE Konferenz (47), 2006.
- [3] P. Francois og O. Bonaventure. Avoiding transient loops during igp convergence in ip networks. 2002.
- [4] B. Halabi. *Internet-Routing-Architekturen*. Hanser Verlag, 1998.
- [5] C. Huitema. *Routing im Internet*. Prentice Hall, 1996.
- [6] A. S. Tanenbaum. *Computernetzwerke*. Pearson Studium, 2000.
- [7] Z. Zhong, R. Keralapura, S. Nelakuditi, Y. Yu, J. Wang, C.-N. Chuah og S. Lee. Avoiding transient loops through interface-specific forwarding, 2005.