Anforderungen effizient umsetzen

Netzplanung von Local Area Networks

Kai-Oliver Detken

Den Kampf um die Technik im LAN-Bereich hat längst Ethernet für sich entschieden. Es sind kaum noch Netze vorhanden, die nicht auf dieser Technik basieren oder nicht zukünftig von ihr abgelöst werden. Jedoch wird auch heute in den meisten Fällen das lokale Netz nur nach Datenapplikationskriterien aufgebaut und neue Leistungsmerkmale heutiger 1/10-Gigabit-Ethernet-Netze zu wenig ausgenutzt.

Dr.-Ing. Kai-Oliver Detken arbeitet als freier Autor im IT-Umfeld und ist Geschäftsführer der Decoit GmbH in Bremen

Das Backbone-Netz eines Unternehmens ist heute die Basis für alle Dienste und Applikationen, die innerhalb sämtlicher Arbeitsprozesse bereitgestellt werden müssen. Das beinhaltet auch, daß die Ausfallzeiten je nach Anforderung so gering wie möglich sein müssen.

Während das Ur-Ethernet keinerlei Mechanismen für die redundante Auslegung des Netzes bereithielt, bietet das Gigabit-Ethernet dafür verschiedene Merkmale:

- lokaler und Inter-Switch Virtual LAN (VLAN) Support;
- Non-Blocking Switching Fabric mit Wire-Speed Switch Performance;
- redundante Netzteile;
- Link Aggregation zur Kapazitätserhöhung einzelner Trunks;
- Meshing-Verfahren auf Layer-2-Ebene;
- Port Mirroring zur Überwachung des Datenverkehrs.

Um eine vollständige Redundanz des Netzes umzusetzen, müssen auf Layer-2- und Layer-3-Ebene unterschiedliche Protokolle unterstützt werden:

- Layer 2: Rapid Spanning Tree Protocol (RSTP);
- Layer 3: Virtual Redundant Router Protocol (VRRP).

Das Spanning Tree Protocol (STP) ist eine 802.1D-Spezifikation, die von der IEEE definiert wurde und stellt einen Bridge-basierten Mechanismus dar, um Verbindungsfehler in Netzen zu minimieren. Es erlaubt die Implementierung paralleler Pfade im Netzverkehr. Das bedeutet, daß redundante Pfade abgeschaltet werden, wenn die Hauptpfade funktionstüchtig sind, oder bereitgestellt werden, wenn der Hauptpfad ausfällt.

RSTP hingegen ist ein standardisiertes Verfahren nach 802.1W mit einer wesentlich geringeren Reaktionszeit. Geht man beim STP noch von mehrstelligen Erholungszeiten im Sekundenbereich aus, so reduzieren sich diese Zeiten bei RSTP auf wenige Millisekunden.

Der Switch kann in mehrfache virtuelle Bridges aufgeteilt werden. Jede virtuelle Bridge kann eine unabhängige STP-Instanz starten. Jede STP-Instanz wird Spanning Tree Domain (STPD) genannt. Jede STPD wiederum hat ihre eigene Root-Bridge und ihren aktiven Pfad. Ist sie einmal geschaffen, können ihr ein VLAN oder mehrere VLANs zugewiesen werden. Ein Port kann dabei nur zu einer STPD gehören. Wenn ein Port ein Mitglied eines mehrfachen VLANs ist, dann müssen alle VLANs zur gleichen STPD gehören. Beim Konfigurieren von VLANs und STP ist zu beachten:

- Jedes virtuelle LAN bildet eine unabhängige Broadcast-Domain.
- STP blockiert Pfade, um eine schleifenfreie Umgebung zu schaffen.
- Wenn STP einen Pfad blockiert, können auf dem blockierten Port keine Daten übertragen oder erhalten werden.
- Innerhalb einer gegebenen STPD benutzen alle VLANs den gleichen Spanning Tree.

Es muß sichergestellt werden, daß Multi-STPD-Instanzen eines einzelnen Switches sich einander nicht in der gleichen Broadcast-Domain sehen.

Das Thema in Kürze

In den letzten Jahren wurden starke Anstrengungen unternommen, um das Ethernet endlich echtzeitfähig zu machen, damit es neue Applikationen wie beispielsweise Voice over IP (VoIP) besser unterstützen kann. Gleichzeitig müssen in den heutigen Netzen Redundanzen eingerichtet werden, um die Ausfallzeiten so gering wie möglich zu halten. Dieser Artikel klärt, inwieweit dies mit Ethernet heute zu realisieren ist.

Dies könnte passieren, wenn beispielsweise noch eine externe Bridge benutzt wird, um VLANs zu verbinden, die zu separaten STPDs gehören. Wird eine STPD gelöscht, werden die beteiligten VLANs auch gelöscht. Aus diesem Grund müssen alle VLANs vor dem Löschen der STPD umgestellt werden.

Werden keine VLANs für den Protokoll-Filter „any" auf einem besonderen Port konfiguriert, können die Datenpakete beim Abschalten des Spanning Tree Protocol innerhalb eines VLAN nicht weitergeleitet werden. Verwendet man das STP, um auf diesem Porttyp zu arbeiten, so muß dieses Protokoll auf dem dazugehörigen VLAN eingestellt werden, damit es teilnehmen kann.

Werden VLANs einer STPD zugewiesen, muß man die STP-Konfiguration und die Wirkung auf die Weiterschaltung des VLAN-Datenverkehrs beachten. Das in *Bild 1* illustrierte Netz benutzt das VLAN, um Stammverbindungen zusammenzufassen. Die folgenden fünf VLANs könnte man dafür wie folgt definieren:

- VLAN1: Vertrieb wird auf Switch A, B und M konfiguriert;
- VLAN2: Personal wird auf Switch A, B und M konfiguriert;
- VLAN3: Produktion wird auf Switch Y, Z und M konfiguriert;
- VLAN4: Entwicklung wird auf Switch Y, Z und M konfiguriert;
- VLAN5: Marketing wird auf allen Switches definiert.

Zusätzlich werden zwei STPDs definiert:

- STPD1 enthält die VLANs Vertrieb und Personal;
- STPD2 enthält die VLANs Produktion und Entwicklung.

Das VLAN Marketing gehört zur vorgegebenen STPD, ist aber nicht STPD1 oder STPD2 zugewiesen. Wenn die Switche in dieser Konfiguration starten, konfiguriert das Spanning Tree Protocol jede STPD, so daß es sich keine aktiven Schleifen in der Topologie ergeben. STP kann die Topologie auf verschiedenen Wegen schleifenfrei konfigurieren.

In Layer-3-Netzen mit redundanten Routern müssen Host-Systeme den Status der Router und ihrer Schnittstellen kennen. IP-Hosts verwenden Default-Gateways für die Kommunikation mit anderen Netzen, Subnetzen oder VLANs (*Bild 2*). Bei Fehlern

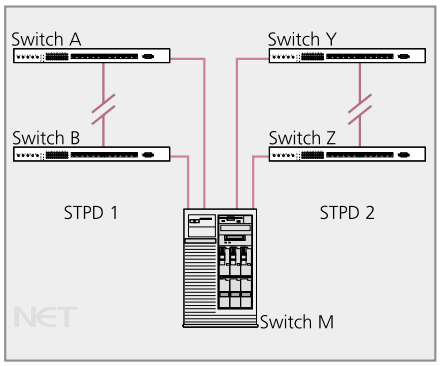


Bild 1: Multiple Spanning Tree Domains

solcher Default-Gateways werden die Hosts normalerweise von dem Netz abgeschnitten. Aus diesem Grund ist der Einsatz von redundanten Gateways zu empfehlen. Dabei muß das Netz in der Lage sein, im Fehlerfall auf das redundante Gateway zu schwenken.

VRRP ist ein IETF-Standard nach RFC-2338-Standard, der es einer Router-Gruppe ermöglicht, als ein einzelnes, virtuelles Default-Gateway aufzutreten. Der Standard beschreibt somit ein Wahlprotokoll, das die Funktion eines Virtual Router liefert. Alle Protokollnachrichten nutzen IP-Multicast-Pakete, obwohl das Protokoll über verschiedene LAN-Techniken arbeiten kann. Jeder virtuelle VRRP-Router besitzt eine einzige bekannte MAC-Adresse. Diese wird als Quelladresse in jeder VRRP-Nachricht durch den Master-Router verwendet, der Bridge Learning in einem LAN ermöglicht.

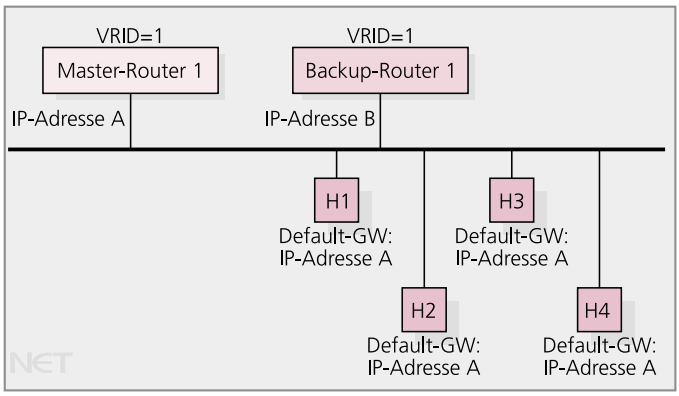
Ein virtueller Router wird durch die Virtual Router Identifier (VRID) und einen Satz von IP-Adressen gekennzeichnet. VRRP-Router können sich mit virtuellen Routern mit ihren realen Adressen auf ihrem Interface verbinden. Ebenfalls können sie mit zusätzlichen Mappings und Prioritäten für Backup-Virtual-Router konfiguriert werden. Das Mapping zwischen VRID und Adressen muß über alle VRRP-Router auf einem LAN koordiniert werden. Allerdings gibt es keine Limitierung gegen eine Doppelnutzung eines Virtual Router Identifier mit verschiedenen Adreßanpassungen auf verschiedenen LANs. Der Gültigkeitsbereich eines jeden virtuellen Routers ist aber begrenzt auf ein einzelnes LAN.

Um den Netzverkehr zu begrenzen, sendet nur der Master für jeden virtuellen Router in periodischen Abständen Advertisement Messages. Ein Backup-Router wird nicht versuchen, dem Master zuvorzukommen, solange dieser eine höhere Priorität besitzt. Dies verhindert Dienstunterbrechungen, solange ein mehr bevorzugter Pfad verfügbar ist. Die einzige Ausnahme ist, wenn ein VRRP-Router immer zu einem Master aller virtuellen Router in Verbindung mit seinen eigenen Adressen wird ist.

Wenn der Master nicht mehr zu erreichen ist, wird der Backup-Router mit der höchsten Priorität nach einer kurzen Verzögerungszeit zum Master. Dieser liefert dann den kontrollierten Übergang der Virtual-Router-Zuständigkeiten mit einer minimalen Unterbrechung aller Services.

VRRP definiert drei Arten der Authentifizierung, um einen einfachen Einsatz in unsicherer Umgebung zu gewährleisten. Es wurde ein Schutz gegen Konfigurationsfehler sowie eine

Bild 2: VRRP-Szenario mit Hilfe von einem virtuellen Router



starke Sender-Authentifizierung in einer sicherheitsbewußten Umgebung implementiert. Eine Analyse möglicher Sicherheitslücken ist in dem Standard enthalten. Zusätzlich können neue Authentisierungsarten jederzeit neu definiert werden, ohne einen Einfluß auf das bestehende Proto-kollformat zu nehmen.

Das VRRP-Protokoll ist so entwickelt worden, daß es eine schnelle Verbindung (1 bis 2 s) vom Backup- zum Master-Router herstellt, um Netzunterbrechungen so gering wie möglich zu halten. Ebenfalls wurde die Protokollkomplexität begrenzt, damit in verschiedenen Szenarien die Master-Kontrolle umgesetzt werden kann. Dies führt zu minimalen Laufzeitanforderungen, aktiven Stati sowie einzelnen Nachrichtentypen und Sendern.

Ein typisches Szenario wäre, wenn zwei redundante Router und/oder deutliche Pfadbevorzugung gegenüber jedem Router vorliegen. Ein Seiteneffekt dieser Annahme wäre, daß duplizierte Pakete in einer kurzen Periode innerhalb der Masterwahl auftreten können. Allerdings tritt dieser Fall sehr selten ein und wird in wenigen Sekunden (hängt von der Netzkonvergenz ab) kompensiert.

Echtzeitfähigkeit implementieren

Um Applikationen wie VoIP zu unterstützen, sollte das eigene LAN CoS-Mechanismen (Class of Service) einsetzen.

Unter CoS wird die Zusammenfassung von gleichartigen Datenströmen zu einer gemeinsamen Klasse verstanden, die dann eine gleichgeartete Dienstgüte vom Netz enthält. Zwei Merkmale muß ein Layer-3-Switch dabei unterstützen:

- Layer 2: IEEE 802.1p;
- Layer 3: Differentiated Service Code Point (DSCP).

Gruppierungen	LAN-Eintritt	LAN-Austritt
IP-QoS	• IP-Sendeadresse • TCP/UDP/andere Ports (Sender oder Empfänger) • IP-Empfangsadresse	• IP-Empfangsadresse • IP Source • TCP/UDP/andere Ports (Sender oder Empfänger)
Destination Address MAC-based	• statisch • dynamisch • Broadcast/unknown Datenratengrenze	• statisch • dynamisch • Broadcast/unknown Datenratengrenze
Paketpriorisierung	• IEEE 802.1p(D)	• Differentiated Service Code Point (DSCP)
physikalisch/logisch	• Source Port • VLAN	• VLAN

Verkehrsgruppierung der CoS-Modi

CoS wird somit als Kontrollmechanismus gesehen, der heterogene Verkehrsmuster bändigen soll. Dadurch wird es möglich, einen Dienst festzulegen, der für eine bestimmte Verkehrsart gilt. Der Hauptvorteil beim Einsatz dieser Art von Class of Service liegt darin, daß man die Kontrolle über eine Verkehrsart besitzt, die eine bestimmte Dienstanforderung von einem System empfängt. Beispielsweise würde ein Videostrom eine höhere Priorität benötigen als ein normaler Datenstrom. Mit Hilfe der CoS-Funktionalität können verschiedene CoS-Profile unterschiedlichen VLANs zugeordnet werden, die Videodaten transportieren.

Folgende Profile können dabei in Blöcken zusammengefaßt werden:

- CoS-Profil: definiert die Bandbreite und Priorisierungsparameter;
- Verkehrsgruppe: Methode zur Klassifizierung oder Gruppierung von Verkehr, der ein oder mehrere allgemeine Attribute besitzt;
- CoS-Policy: Kombination der Ergebnisse von der Einteilung eines CoS-Profils zu einer Verkehrsgruppierung.

CoS-Profile werden eingeteilt in Verkehrsgruppierungen, um das Forwarding-Verhalten des Switch zu modifizieren. Bei Einteilung zu einer Verkehrsgruppe, umfaßt die Kombination der Gruppe zu einem CoS-Profil eine einzelne Policy, die Bestandteil eines Policy-basierten Class of Service ist.

In Abhängigkeit von der Zielsetzung kann ein Teil der CoS-Parameter hervorgehoben und separat betrachtet werden. Dabei müssen nicht immer alle Parameter berücksichtigt werden. Die Verkehrsgruppierung kann in folgende Kategorien unterteilt werden:

- IP-Information (IP-CoS-Gruppierung);
- Destination MAC (MAC-CoS-Gruppierung);
- Packet Priority Information (wie IEEE 802.1p)
- physische/logische Konfiguration (physischer Quellport oder VLAN-Zuordnung).

Ein CoS-Profil ist einer gewünschten Verkehrsgruppe zugeordnet, um eine CoS-Policy zu schaffen. Bei einem Ereignis, in dem ein Paket zu mehreren Gruppenkriterien paßt, gibt es einen vorbestimmten Vorrang (Precedence), um die Verkehrsgruppe zu bestimmen. Allgemein wird der spezielleren Verkehrsgruppe der Vorrang erteilt. Mögliche Verkehrsgruppen und ihre Optionen zeigt die *Tabelle*. Die Gruppierung ist nach der Reihenfolge des Vorrangs (höchster Rang zu niedrigstem) aufgelistet.

Es kann ein Satz von IP-Empfangsadressen für eine IP-CoS-Verkehrsgruppierung durch eine spezielle Netzadresse und Subnetzmaske verwendet werden. Eine IP-CoS-Verkehrsgruppierung kann auch optional andere Komponenten von IP-Paketen enthalten, wie IP-Sende- und Empfangsadresse oder Quell-TCP/UDP-Portinformationen.

Fazit

Um eine ausreichende Redundanz und damit hohe Ausfallsicherheit zu schaffen, genügt es heute nicht Module zweifach auszulegen. Auch das Netz selbst muß hochgradig vermascht werden. Dadurch müssen neue Protokolle zum Einsatz kommen, die es Ethernet ermöglichen, redundante Wege zu nutzen. Ein wichtiger Schritt stellt VRRP dar, das zum ersten Mal standardkonform eine Lösung anbietet, was bisher immer proprietär durch die Hersteller gelöst wurde. Als

Alternative zu VRRP kann zwar der Einsatz eines dynamischen Routing-Protokolls (z.B. OSPF) in Erwägung gezogen werden, dies erzeugt aber einen nicht unwesentlich höheren Aufwand und muß situationsbedingt abgewogen werden.

Desweiteren kann heute nicht mehr auf das Spanning-Tree-Protokoll verzichtet werden. Es ist für die Redundanz zuständig, indem passive Pfade abgeschaltet und erst aktiv werden, wenn der Switch oder die Verbindung ausfallen. Durch die Weiterentwicklung des Standards IEEE 802.1W ist es auch für heutige Anforderungen bestens gerüstet. VRRP ist für die Redundanz ausschließlich auf Layer 3 zuständig und wird erst wirksam, wenn das Default-Gateway Probleme bekommt. Beide Protokolle können dabei problemlos zusammen eingesetzt werden.

Mit Class of Service werden die Probleme im Netz, die bei hoher Auslastung und Übergängen mit unterschiedlichen Bandbreiten entstehen, beseitigt. Router und Switche reihen bei Überlastung die eingehenden Pakete in Warteschlangen ein und schicken diese dann früher oder später weiter. Zur Flußsteuerung beim Filetransfer einigen sich die Transportschichten der beiden Kommunikationspartner in solchen Fällen auf eine niedrigere Transferrate.

Für Echtzeitapplikationen ist dieses Verfahren jedoch ungeeignet, da die auftretenden Verzögerungen zu Störungen der Übertragungsqualität führen. Durch die Charakterisierung gemäß einer CoS, können die Priorisierungsmechanismen in TCP/IP-Netzen gezielt zur Festlegung garantierter Bandbreiten genutzt werden.

Die Priorisierung setzt dabei auf verschiedenen Ebenen an. So können hochwertige Router oder Layer-3/4-Switche, die in den Datenpaketen vorhandenen Informationen auswerten und die Pakete in Abhängigkeit vom Inhalt unterschiedlich behandeln. Die Unterscheidung wird dabei nach Protokoll oder IP-Adressen getroffen. Moderne Layer-4-Switche identifizieren zusätzlich die individuellen Applikationen über die TCP/UDP-Portnummern und behandeln die Pakete ggf. bevorzugt. Ein weiterer Service dieser Komponenten ist die Möglichkeit, ausgehende Pakete entsprechend ihrer Priorität zu kennzeichnen, um damit die fehlenden Funktion der weniger intelligenten Komponenten zu ergänzen. Durch geeignete Header mit Differentiated Services Code Point (DSCP) oder angefügte VLAN-Tags (IEEE 802.1p) werden die Datenströme gekennzeichnet.

Beim Einsatz von CoS ist mit Problemen hinsichtlich der alten Netzkomponenten zu rechnen, wenn diese das neue Frame-Format nicht verarbeiten können. Daher müssen neu anzuschaffende Geräte unbedingt den Standard IEEE-802.1p/Q unterstützen. Am Übergang zwischen LAN-WAN endet eine Paketpriorisierung auf Layer 2 und muß zwingenderweise auf Layer 3 weitergeführt werden. Die verschiedenen Router-Hersteller arbeiten an einer Implementierung, um hier einen bidirektionalen LAN-WAN-Übergang zu schaffen. *(bk)*