

6.2.7 スイッチの冗長化

スイッチ間の経路を冗長化するため、STP (Spanning Tree Protocol) を用いる。

なお、STP は、スイッチのことを「ブリッジ」と称し、ブリッジの名を冠した様々な用語を定義している。STP が規格化された当時は「ブリッジ」と呼ばれていたからだ。本書もそれに合わせ、以降の解説では「ブリッジ」を用いることにする。

● STP の技術動向

詳しくは後述するが、STP を用いると、経路の切替えに最大で 50 秒かかる。これに対し、リンクアグリゲーションとスタック接続を用いると、経路の切替えは瞬時に行われる。

STP はアクティブ／スタンバイ型なので、通常時に未使用のリンクが生じる。これに対し、リンクアグリゲーションはアクティブ／アクティブ型なので、通常時に全てのリンクの帯域を活用できる。

このように、今日では STP より優れた技術が普及している。こうした背景から、現在、STP は徐々に使われなくなっている。

試験では、過去にしばしば出題されており、近年でも時折登場している。そこで、補遺の扱いであるが、STP をここで解説する次第である。

● スパニングツリーの仕組み

ブロードキャストドメイン内で STP を用いると、一つのブリッジを頂点（ルート）とする、ツリー構造のトポロジーが形成される。

このツリー構造を構成するため、特定のブリッジの特定のポートが論理的にブロックされる。通常のフレームは、ブロックされたポートを経由して転送されることがない。

ツリー構造のルートに位置するブリッジを**ルートブリッジ**という。

このツリーのことを**スパニングツリー**という。これは、スパニングツリーアルゴリズムに基づいて計算されたものである。

障害が発生したとき、新たなツリー構造が再構成される。

詳説

STP の基本的な機能について、第 6 章の「6.2.1 リンクの冗長化」の「● STP」で解説しているので、参照していただきたい



試験に出る

スパニングツリーはしばしば出題されている。令和 3 年午後 II 問 1、平成 30 年午後 I 問 2、平成 26 年午前 II 問 5、平成 25 年午後 II 問 1、平成 24 年午後 II 問 2、平成 22 年午後 II 問 2、平成 21 年午後 I 問 1、平成 21 年午後 II 問 2 で出題された



IEEE802.1Dは、MACブリッジに関する標準であり、スパニングツリー以外にも様々な規定がなされている



ブリッジ内の全ポートのMACアドレスのうち、最小の値を「xx-xx-xx-xx-xx-xx」とすると、工場出荷時のブリッジプライオリティは「0x8000」なので、ブリッジIDは「80-00-xx-xx-xx-xx-xx」となる。ただし、このままでは、セグメント内のルートブリッジは、MACアドレスの値によって決定されることとなるため、上位2バイトのブリッジプライオリティを変更して、ブリッジIDの優先順位を明示的に決定するのが一般的である

スパニングツリーの構成、及び、障害発生時の再構成のため、ブリッジ間で定期的に情報交換が行われる。このときやり取りされるフレームを **BPDU** (Bridge Protocol Data Unit) という。

● **スパニングツリーを構成する手順**

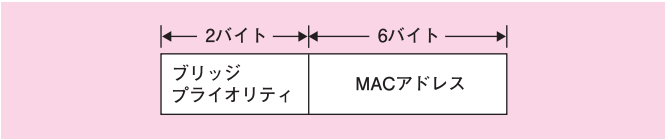
STPを動作させているネットワークでは、ブリッジID、 **パスコスト**、ポートIDに基づいて特定のポートを論理的にブロックすることで、スパニングツリーを構成する。

具体的に言うと、以下に示す手順に従って、スパニングツリーが構成される。

- 1. ルートブリッジの決定
- 2. ルートポート、代表ポート、オルタネートポートの決定
- 3. スパニングツリーの維持と障害検出

まず、スパニングツリーのルートノードに当たる **ルートブリッジ** を決定する。これはブリッジIDによって決定され、ブリッジIDの値が最も小さいブリッジがルートブリッジとなる。

ブリッジIDのフォーマットを次に示す。



図：ブリッジID

それぞれの用語の意味を次に示す。

- **ブリッジプライオリティ**
デフォルトは 32768 (0x8000)。設定により任意の値に変更できる。
- **MAC アドレス**
ブリッジがポートごとにもつ MAC アドレスの中から一つ選ばれる。通常、最も小さい値が選ばれる。

ルートブリッジは、パスコストを 0 に設定した BPDU を、すべてのポートから接続先のブリッジに向けて送信する。次いで、

BPDUを受信したブリッジは、設定されたパスコストを加算したBPDUを、受信したポート以外のポートから送信する。以降、接続先のブリッジは同様に振る舞い、BPDUが同一ブロードキャストドメインの全域に行き渡る。

各ブリッジは、ルートブリッジを起点に伝搬されたBPDUに基づき、ルートブリッジへのコスト（ルートブリッジから当該ブリッジに至るまでの経路のパスコストを加算した値）が分かっている。これを踏まえ、各ブリッジは、各ポートの役割を決定する。

その役割は、**ルートポート** (root port)、**代表ポート** (designated port)、**オルタネートポート** (alternate port) の3種類がある。パスコスト、ブリッジID、ポートIDに基づいてこの三つのうちのいずれか一つになる。

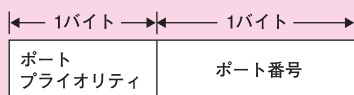
表：スパニングツリーを構成するブリッジのポートの役割

| 役割 | 説明 |
|-----------|---|
| ルートポート | ブリッジの中で、ルートブリッジへのコストが最小のポート。ルートポートの候補が複数ある場合は、送信元ブリッジ*IDの小さい方、(同一ブリッジの場合は) 送信元ポート*IDの小さい方が選出される。ただし、ルートブリッジはルートポートをもたない |
| 代表ポート | 指定ポートともいう。ブリッジを収容しているセグメントの中で、ルートブリッジへのコストが最小のポート。代表ポートをもつブリッジは、そのセグメントの 代表ブリッジ となる。代表ポートの候補が複数ある場合は、送信元ブリッジ*IDの小さい方、(同一ブリッジの場合は) 送信元ポート*IDの小さい方が選出される |
| オルタネートポート | ルートポートでもなく、代表ポートでもないポート。 ブロッキングポート ともいう。論理的にブロックされた状態になり、データフレームは送受信しない。BPDUは送信しないが、受信だけ行う |

※「送信元ブリッジ」とは、ルートブリッジから自分に至る経路の中で、自分と接続しているブリッジ（自分の直前にあるブリッジ）を指す。

※「送信元ポート」とは、送信元ブリッジの中で、自分と接続しているポートを指す。

ポートIDのフォーマットを次に示す。



図：ポートID

詳説

代表ブリッジの定義から明白のように、ルートブリッジは自分を収容しているセグメントにおける代表ブリッジとなる

詳説

IEEE802.1D規格は「オルタネートポート」と表記している。とはいえ、一般的にはブロッキングポートと呼ばれることも多い。試験ではブロッキングポート、非指定ポートと表記されることがある

詳説

リンクの帯域に基づいてパスコストの範囲と推奨値が定められている。10Mbpsは範囲が「50～600」で推奨値が「100」。100Mbpsは範囲が「10～60」で推奨値が「19」。1Gbpsは範囲が「3～10」で推奨値が「4」である



試験に出る

平成6年午前II問5では、ネットワークのバスコストに基づいてツリー構成が問われた



詳説

Hello時間(BPDUをやり取りする間隔)は、デフォルトで2秒。最大エージ時間(BPDUを受信できない期間の許容値)は、デフォルトで20秒。転送遅延時間(リスニング状態からラーニング状態、ラーニング状態から転送状態へ遷移するまでの待ち時間)は、デフォルトで15秒である。なお、スパンニングツリーの再計算は、ブリッジ単位で行われる。障害発生や構成変更の影響を受けていないブリッジでは再計算は行われない

それぞれの用語の意味を次に示す。

- ポートプライオリティ

デフォルトは 128 (10 進数表記)。設定により任意の値に変更できる。

- ポート番号

ポートの物理的な番号。

● スパニングツリーの構成例

スパニングツリーの構成例を以下に示す。

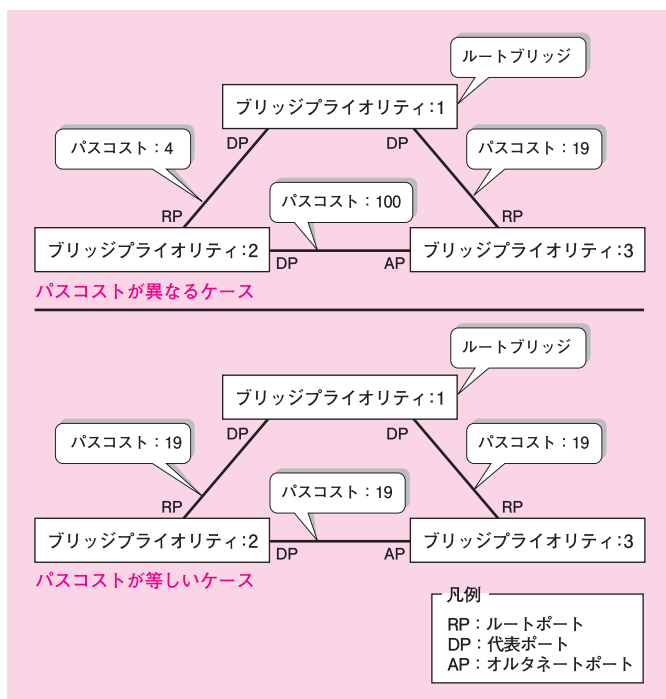


図: スパニングツリーの構成例

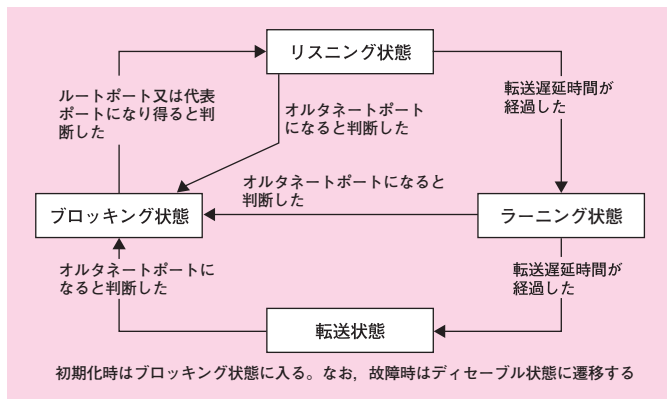
● ポートの状態遷移

スパニングツリーが完成した後は、ブリッジはBPDUをやり取りしてスパニングツリーを維持する。直接接続されたリンクのダウンを検出した場合や、最大エージ時間が経過してもBPDUを受信できなくなった場合、障害が発生したと判断し、スパニングツリーの再構成が行われる。

詳説

さらには、直接接続されたリンクのアップを検出した場合や、ルートポート以外のポートで、ルートブリッジまでのパスコストがより小さいBPDUを受信した場合もネットワークの構成が変更されたと判断し、スパンニングツリーの再構成が行われる。

このとき、ポートは次に示す状態遷移を行う。



図：ポートの状態遷移

● ブロッキング状態

オルタネートポートは、ブロッキング状態になっている。データフレームは転送しない。BPDUは送信しないが受信だけは行い、BPDU受信の結果、ルートポート又は代表ポートになり得ると判断したなら、リスニング状態に遷移する。

● リスニング状態

転送遅延時間の期間中、スパンニングツリーの計算を行う。データフレームは転送しない。BPDUの送受信を行う。オルタネートポートになると判断したらブロッキング状態に遷移する。さもなくば、ラーニング状態に遷移する。

● ラーニング状態

転送遅延時間の期間中、再構成されたスパンニングツリーのトポロジに合わせて、MACアドレステーブルの学習を行う。データフレームは転送しない。BPDUの送受信を行う。

● 転送状態

ルートポート又は代表ポートのどちらかの状態である。データフレームを転送し、BPDUの送受信を行う。

物理的にトポロジが変化すると、この変化の影響を受けるブリッジはSTPの再計算を行い、スパンニングツリーを再構成する仕組みになっている。

それだけでなく、変化を検知したブリッジは、ルートブリッジに対し、変化した旨を通知する。そして、ルートブリッジから全てのブリッジに対し、エージングタイムを15秒にセットする旨、通知する仕組みになっている。このエージングタイムの短縮は、スパンニングツリーを再構成している間の35秒間(ポート状態がブロッキングとリスニングに留め置かれる期間に相当)にのみ生じる、一時的なものである。

エージングタイムを短縮する目的は、STPの再計算を行っているブリッジのMACアドレステーブルをいったんクリアし、新たに形成される論理的トポロジに従ってMACアドレステーブルのエントリを再生成するためである。ポート状態がラーニングに遷移すると、ブリッジはMACアドレスを学習し始める。やがて、スパンニングツリーが再構成され、エージングタイムは元の値(300秒)に戻される

詳説

フレーム転送遅延時間、コンバージェンス時間を考慮し、一つのセグメントでブリッジホップは7以下にすることがIEEE802.1Dで規定されている



試験に出る

スパンニングツリーの再構築中は通信ができない。この点について、平成30年午後I問2、平成21年午後I問1で出題された

ブロッキング状態からリスニング状態、ラーニング状態を経て転送状態に遷移し、スパニングツリーが完成することを、**コンバージェンス**（収束）という。この経過時間は、次の式が示すとおり、デフォルトで 50 秒となる（括弧内はデフォルト値）。

$$\text{最大エージ時間 (20 秒)} + \text{転送遅延時間 (15 秒)} + \text{転送遅延時間 (15 秒)} = 50 \text{ 秒}$$

● 様々な STP (RSTP, MSTP)

現在、STP を高速化した **RSTP** (Rapid Spanning Tree Protocol, 高速スパニングツリープロトコル) が IEEE802.1D:2004 (旧 IEEE802.1w) として標準化されている。RSTP のコンバージェンス時間は数秒単位となっている。

RSTP では、STP で定義されていたオルタネートポートの役割がさらに細分化され、**オルタネートポート**（ルートポートのスタンバイ）、**バックアップポート**（代表ポートのスタンバイ）の二つが規格化されている。また、RSTP では、ネットワーク構成の変化を検知したら隣接するブリッジ間でやり取りするようになり、やり取りされる内容もより詳細なものに変更されている。このように様々な改善点が盛り込まれたことにより、切替え時間が短くなっている。

IEEE802.1D で規格化された STP は、VLAN ごとにスパニングツリーを構築することができない。VLAN ごとにスパニングツリーを構築するには、IEEE802.1s で規格化された **MSTP** (Multiple Spanning Tree Protocol) を使用する。



試験に出る

RSTP について、令和3年午後Ⅱ問1で出題された。なお、RSTP の知識がなくても、STP の知識があれば本文中の説明に基づき正解できるように配慮されていた。

MSTP について、平成24年午後Ⅱ問2で出題された