Cisco in Persian

سیسکو به پارسی

Spanning Tree Protocol

نوشته: شفق زندی
http://www.zandi.ir

فهرست

طراحی شبکه لایه دو 4
مفاهیم اصلی در STP 6
مراحل Spanning Tree 7
انتخاب یک Root Bridge 7
چهار معیار اصلی STP در تصمیم گیری ها 8
انتخاب Root Port ها 9
انتخاب Designated Port ها 10
حالات مختلف Port 11
زمان بندی و تایمرها در STP 13
پیغام های STP 14
تغییر Topology در STP 15
STP به ازای هر VLAN 18
تعیین Root Bridge 19
انعطاف پذیری در تنظیمات STP 20
بهبود Convergence در STP 21
PortFast 21
UplinkFast 22
BackboneFast 24
مقابله با BPDU های مزاحم 26
BPDU Skew Detection 27
Loop Guard 27
Unidirectional Link Detection 27
RTSP 29
وضعیت BPDU ها در RSTP 31
انواع پورت در RSTP 32
RSTP Synchronization 33
تغییر توپولوژی در RSTP 34
Multiple Spanning Tree 35
MST Region 36
تنظیمات MST 38

Cisco in Persian

مقدمه ای بر Spanning Tree Protocol

در لایه سه Routing Protocol ها به ازای مسیرهای Down شده، مسیری جدید برای رسیدن به مقصد انتخاب میکنند. در شبکه علاوه بر سرعت انتقال و کارایی بالا مسائلی نظیر رفع خطا و حل مشکلات بصورت داینامیک و پویا حائز اهمیت است. در لایه دو، برای حل مشکلات لینکها و استفاده از مسیرهای Redundant – "مسیرهای افزونه" از STP یا Spanning Tree Protocol که در IEEE 802.1D تعریف شده، استفاده می کنیم. کار STP در یک جمله خلاصه میشود: شبکه لایه دو نباید Loop داشته باشد یعنی برای رسیدن یک فریم به مقصد تنها یک مسیر لایه دو وجود داشته باشد.

چند اصل مهم در رابطه با لایه دو شبکه های ایترنت:

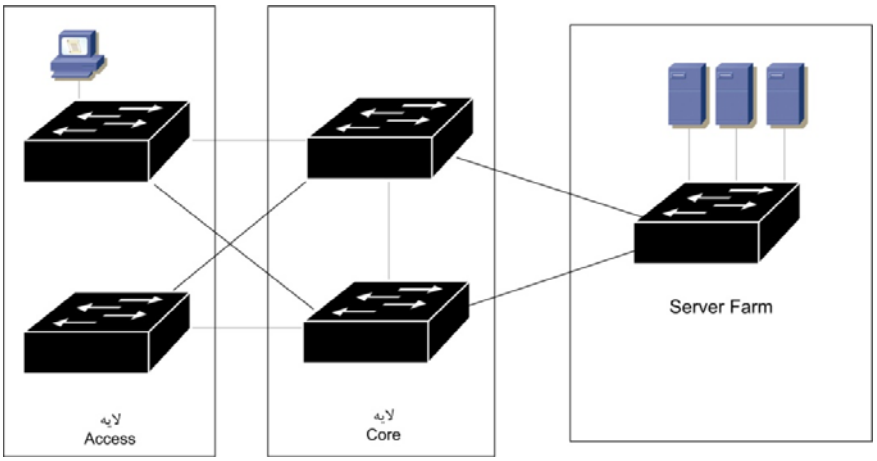
- Bridge در سطح شبکه، Transparent یا شفاف است، هیچ تغییری در Frame ایجاد نمیکند.
- فریم ها بر اساس CAM (Content Addressable Memory) به مقصد خود فرستاده میشوند.
- فریم های Broadcast به همه Port های آن VLAN ارسال میشوند.
- فریم های Unknown Unicast مقصدشان در CAM مشخص نیست به همه پورت ها فرستاده میشوند.
- وقتی فریمی بین دو سوییچ، متناوبا ارسال و دریافت میشود، Bridging Loop رخ داده است.

Cisco in Persian

Spanning Tree Algorithm یک الگوریتم است که براساس اطلاعات دریافتی از سوئیچ های همسایه، یک نقطه مرکزی و Root انتخاب کرده و تمام مسیر ها را تا آن نقطه محاسبه میکند تا شبکه ای-Loop Freeبصورت یک درخت با شاخ و برگهایش بسازد.

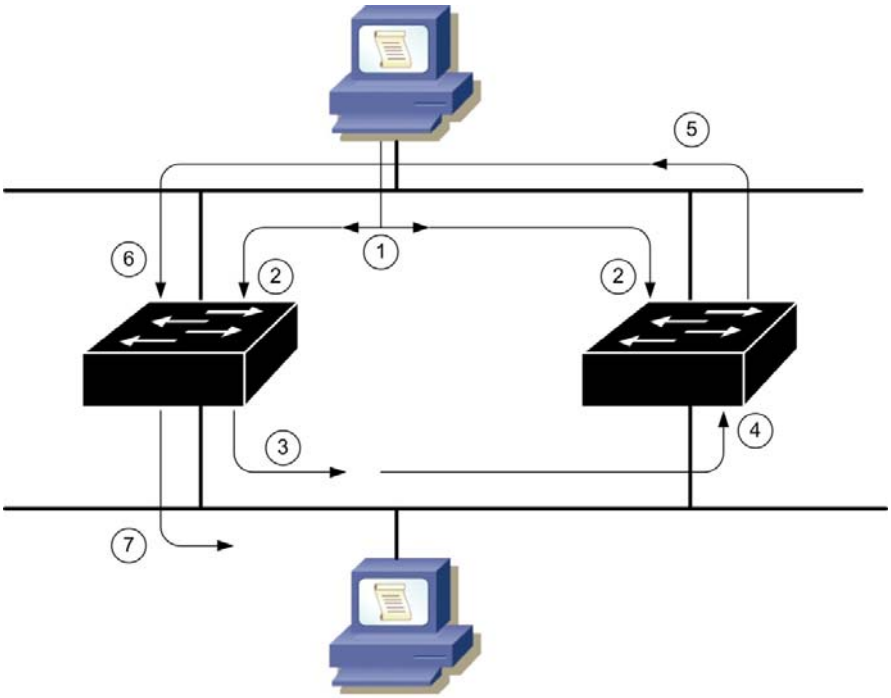
طراحی شبکه لایه دو

شبکه ها باید Loop طراحی شوند! به این دلیل که مسیر های Redundant بوجود بیاید تا در موقع نیاز مسیر دیگری علاوه بر مسیر اصلی تا مقصد وجود داشته باشد. اما باید توجه داشت که درآن واحد و یک زمان در لایه دو، باید یک تنها یک مسیر بدون لوپ، فعال باشد. این مطلب برای لایه سه و IP Routing صدق نمیکند و میتوانید بار شبکه را بین چند Route موازی به یک مقصد پخش کنید.



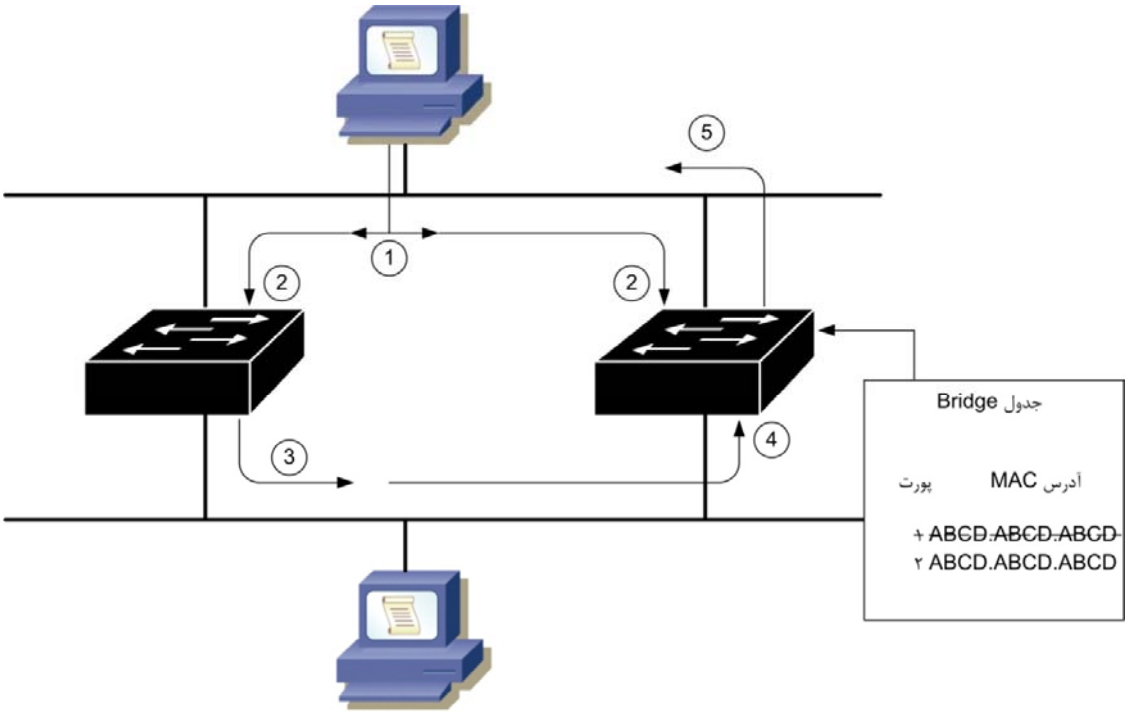
وقوع Loop در شبکه در شکل زیر نشان داده شده است. بدون STP ، Broadcast شبکه شکل زیر را دچار Loop Feedback میکند:

Cisco in Persian



مراحل ارسال فریم تکراری بین دو سوییچ در شبکه بدون STP

در شکل زیر بدونSTP ، Unicast ها نیز شبکه را دچار مشکل کرده و Bridge Table را خراب میکنند:



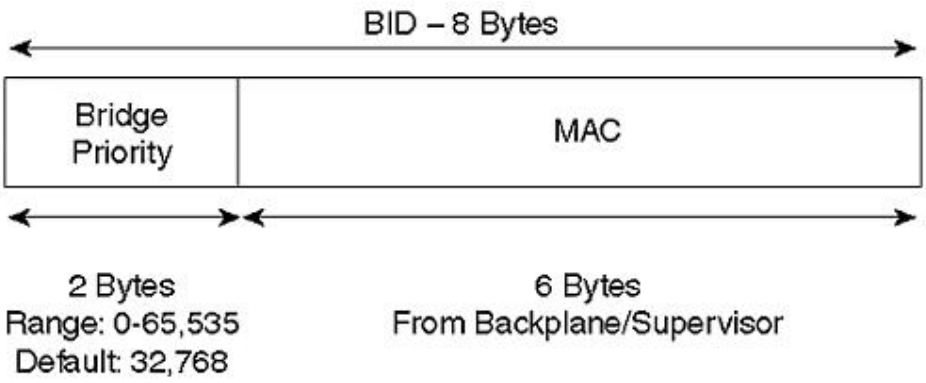
تخریب جدول **MAC** آدرس ها توسط فریم تکراری بین دو سوییچ
در شبکه بدون **STP**

مفاهیم اصلی در STP

همانطور که اشاره شد، برای اینکه شبکه را بصورت یک گراف؛ یک درخت بدون Loop بسازد نیاز به انتخاب یک Root یا سوییچ اصلی که ریشه این شاخ و برگ ها است دارد.

محاسبات STP بر اساس Bridge ID و Path Cost صورت میگیرد.

برای اینکه سوییچ Root را انتخاب کنیم؛ از Bridge ID استفاده میکنیم هرچه کمتر باشد شانس Root شدن سوییچ بیشتر میشود. Bridge ID هشت بایت است و از ترکیب 2 بایت بعنوان اولویت و 6 بایت MAC تشکیل میشود:



در نسخه DEC STP از 8 بیت برای Bridge Priority استفاده میشد که در سوییچ های سیسکو از نسخه IEEE STP با 16 بیت Bridge Priority پشتیبانی میشود که عدد 32768 بصورت پیش فرض است.

Path Cost یا ارزش مسیر مشخص کننده ارزش هر لینک تا سوئیچ کناری است. این اعداد بصورت استاندارد در 802.1D براساس سقف پهنای باند 1000Mbps تنظیم شدند. اما از آنجا که با رشد تکنولوژی و شبکه این عدد برای رسانه های بالا 1Gbps راه حلی نداشت، جدول Cost جدید وضع گردید که دیگر مثل قبل تناسبی نیست.

Bandwidth	STP Cost
4 Mbps	250
10 Mbps	100
16 Mbps	62
45 Mbps	39
100 Mbps	19
155 Mbps	14
622 Mbps	6
1 Gbps	4
10 Gbps	2

مراحل Spanning Tree

در اجرای STP، برای داشتن شبکه ای Loop Free و در خلال هر Convergence این مراحل طی میشوند:

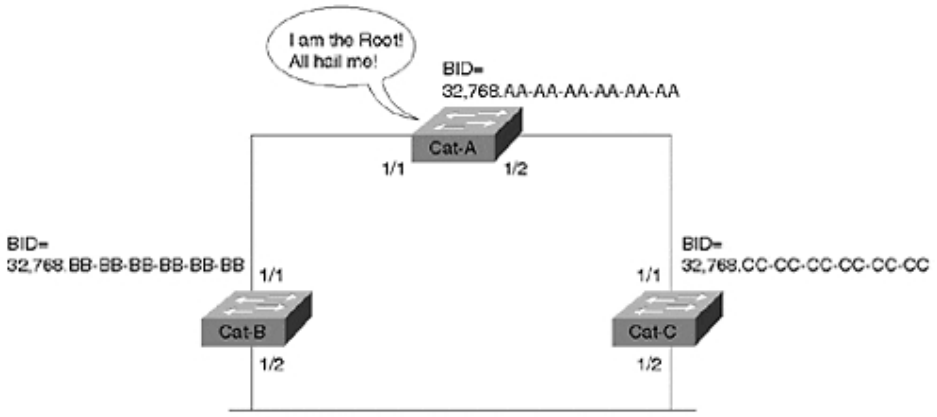
1- انتخاب یک Root Bridge
2- انتخاب Root Port ها
3- انتخاب Designated Port ها

انتخاب یک Root Bridge

نقطه اصلی و ریشه ای که محاسبات STP بر اساس آن صورت میگیرد، Root Bridge است.
Root Bridge دارای کوچکترین BID در سطح شبکه است و از این رو است که بعنوان Root انتخاب شده پس هرگاه سوئیچی با BID کوچکتر به شبکه متصل گردد، آن سوئیچ تبدیل به Root Bridge خواهد شدو محسبات از سرگرفته میشود و لینک های لوپ شده دورتر به نسبت آن حذف میشوند. هر دو ثانیه Root Bridge خود را در شبکه تبلیغ میکند.

Cisco in Persian

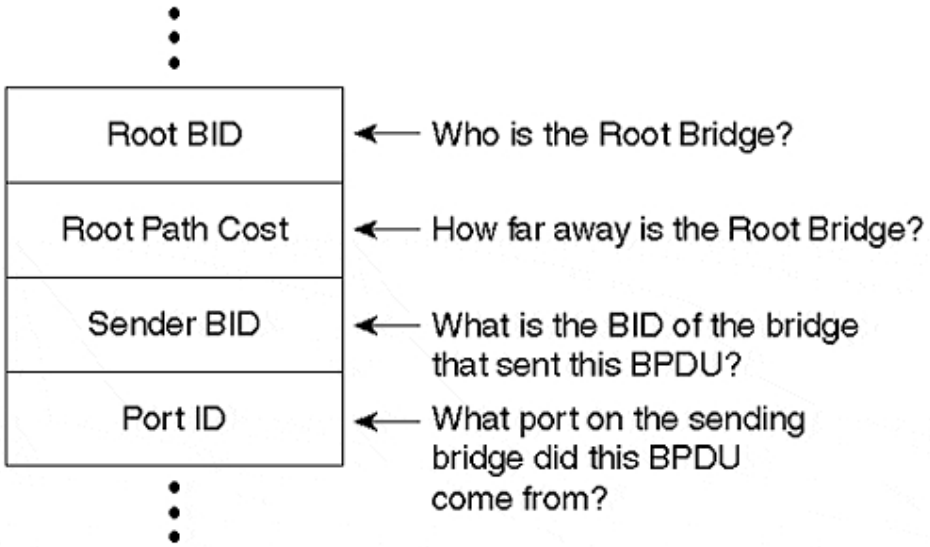
Configuration BPDU حاوی اطلاعات Spanning Tree است که تنها از سوی Root Bridge در شبکه توزیع میگردد و همانطور که اشاره شد، هر دو ثانیه Root Bridge خود را در شبکه تبلیغ میکند. در ابتدا هر سوئیچ خود را Root Bridge دیده و شروع به تبلیغ خود میکند تا زمانی که یک Configuration BPDU از یک Root Bridge با BID بهتر (اولویت بهترکمتر و یا در حالت مساوی بودن اولویت، MAC Address کوچکتر) ببیند این کار را ادامه میدهد. اما پس از پیدا شدن یک سوئیچ با BID بهتر، سوییچ یا سوییچ ها آن را بعنوان Root Bridge به زیر شاخگان خود معرفی میکنند.



چهار معیار اصلی STP در تصمیم گیری ها

در تمام تصمیم گیری های STP، در تشکیل توپولوژی شبکه از اولویتهای زیر به ترتیب برای انتخاب استفاده میشود:

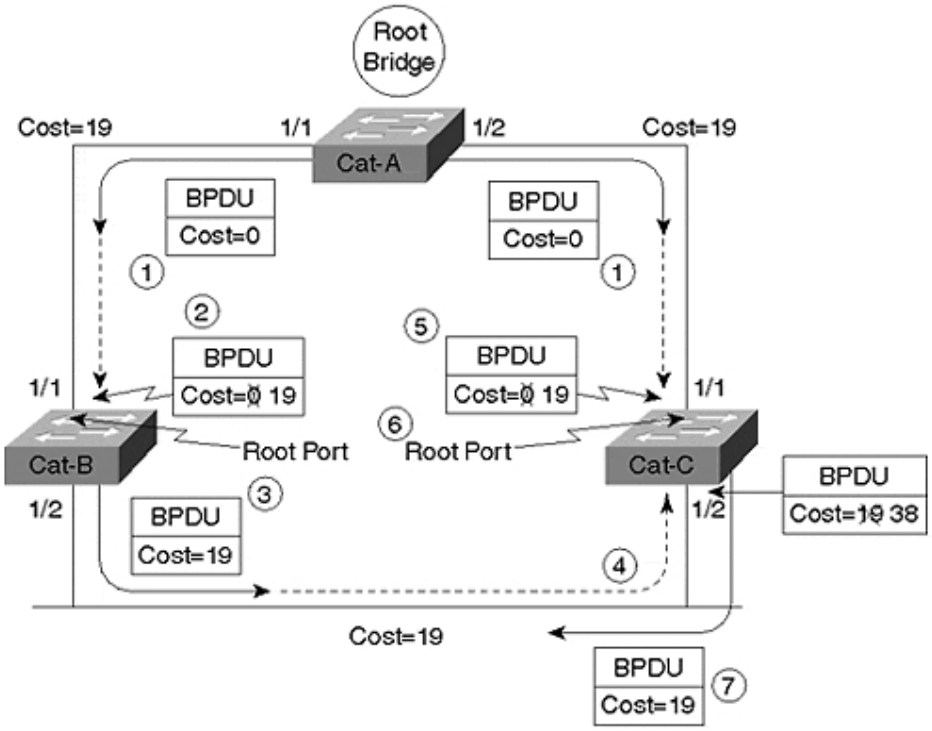
1. Lowest Root BID
2. Lowest Path Cost to Root Bridge
3. Lowest Sender BID
4. Lowest Port ID



انتخاب Root Port ها

وقتی BPDU از یک پورت دریافت میشود، Cost آن محاسبه میگردد. این عمل به سادگی از مجموع عدد Cost روی فریم BPDU بعلاوه Interface Cost حاصل میشود. در واقع Path Cost عددی است که به Interface تعلق دارد و Root Path Cost از اضافه شدن Path Cost به Cost مندرج در Configuration BPDU بدست می آید.

بهترین عدد بدست آمده روی سوییچ، Root Port آن سوییچ میشود یعنی سوییچ Root شبکه را از آن پورت می بیند. انتخاب Root Port برای همه سوئیچ ها به غیر از Root Bridge انجام میگیرد.



تمام سوئیچ ها (غیر از Root Bridge) به یک Root Port نیاز دارند.

باید توجه داشت که محاسبه Cost تنها در زمان دریافت BPDU از پورت صورت میگیرد نه در زمان انتقال و خروج آن از سوییچ.

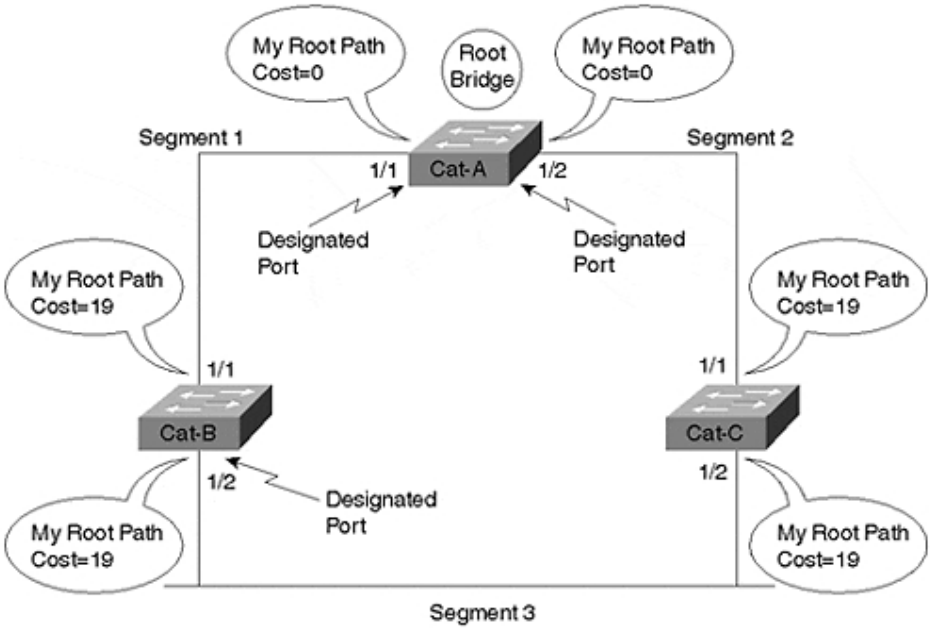
انتخاب Designated Port ها

در صورتیکه که به یک Segment در LAN، دو یا چند سوئیچ متصل باشند، تنها یک سوئیچ باید فریم های Segment را دریافت و ارسال کند. (برای جلوگیری از Loop و نگهداری از Bridging Table)

پورتی که وظیفه اتصال سگمنت به LAN را دارد، Designated Port نامیده میشود و دارای بهترین Root Path Cost است اگر مساوی برقرار شد انتخاب براساس چهار معیار اصلی که قبلا عنوان شد صورت میگیرد.

به ازای هر سگمنت یک Designated Port داریم و به ازای هر سوئیچ یک Root Port.

تمام پورتهای Root Bridge، Designated Port هستند.



در واقع ایده این است که اگر به ازای هر سگمنت تنها یک پورت مامور انتقال داده ها باشد، شبکه دچار Loop نمیشود. سوئیچی را که در سگمنت Designated Port دارد، Designated Bridge مینامیم.

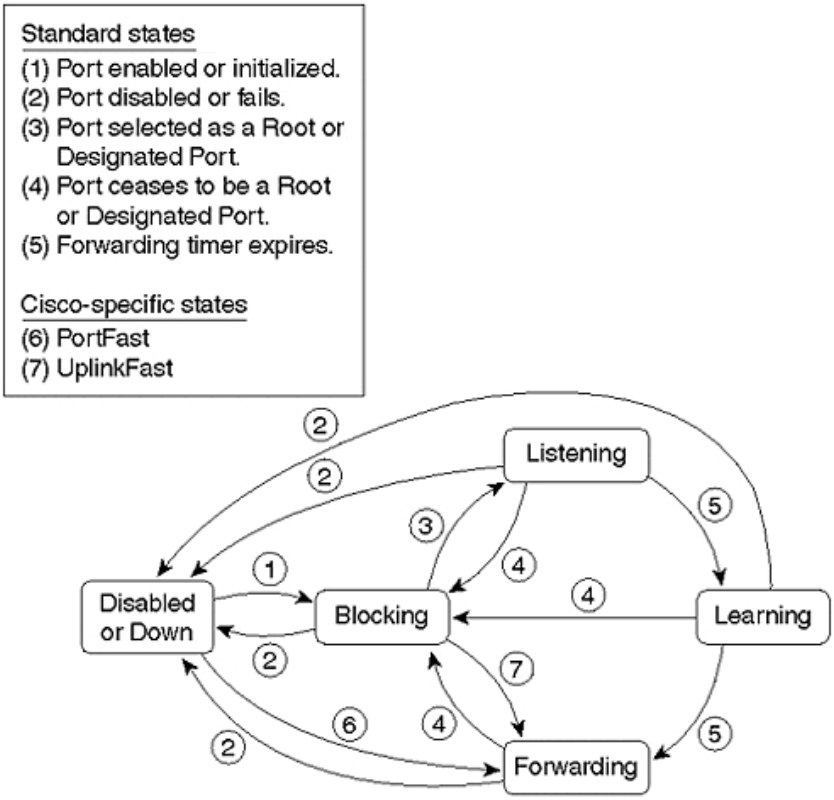
هر پورتی که Root Port یا Designated Port نباشد، به حالت Blocking میرود.

حالات مختلف Port

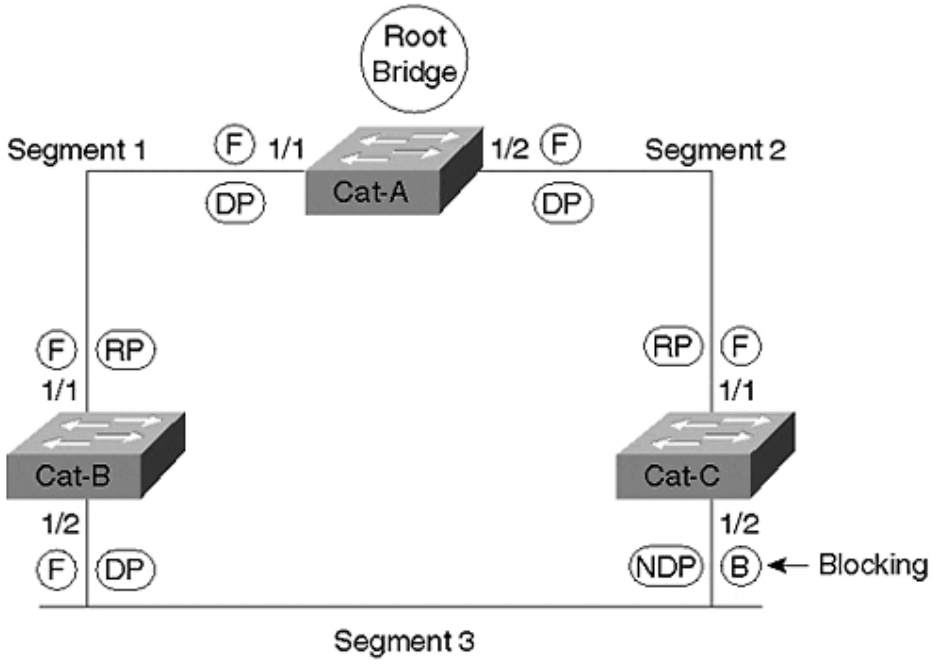
برای اینکه STP درست کار کند، هر پورت مراحل مختلفی را پشت سر میگذارد. پورتها کار خود را از حالت غیر فعال شروع کرده، در چند مد Passive قرار گرفته و نهایتا در صورت اینکه STP اجاز دهد، Active میشوند. جدول زیر از پایین به بالا مراحل را شرح میدهد:

وضعیت	وظیفه	Purpose
Forwarding	انتقال ترافیک را شروع میکند.	Sending/receiving user data
Learning	در حال یادگیری جدول Bridging است.	Building bridging table
Listening	به ارسال BPDU میپردازد.	Building "active" topology
Blocking	به BPDU ها گوش میکند.	Receives BPDUs only
Disabled	پورت غیر فعال است.	Administratively down

بین حالت Listening و Learning زمانی، بنام Forward Delay باید سپری شود.



یک پورت وقتی به حالت forwarding میرود که بتواند Root Port یا Designated Port شود.



مطابق شکل بالا وضعیت هر پورت در کنار آن نوشته شده است.

Cisco in Persian

مشخصه	وضعیت پورت
B	Blocking
F	Forwarding
DP	Designated Port
RP	Root Port
NDP	Non-Designated Port

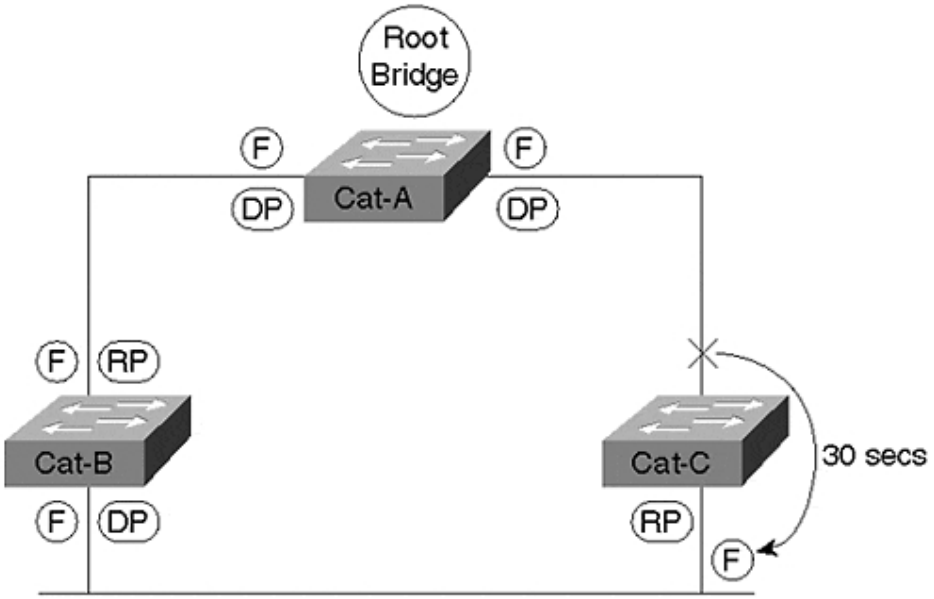
با دستور `show spanning interface` میتوان وضعیت هر interface را چک کرد. همچنین برای debug کردن تغییر وضعیت پورت ها از دستور `debug spanning-tree switch state` میتوان استفاده کرد.

مدت زمان Listening و Learning هر کدام 15 ثانیه بطول می انجامد.

زمان بندی و تایمرها در STP

STP از سه Timer برای کار خود استفاده میکند. این زمان بندی توسط Root Bridge و در پیغام های Configuration BPDU به بقیه اعلام میگردد.

Timer	هدف	Default
Hello Time	زمان ارسال Configuration BPDU توسط Root Bridge	2 ثانیه
Forward Delay	مدت زمان حالات Listening و Learning	15 ثانیه
Max Age	عمر BPDU	20 ثانیه



در شکل بالا Cat-C بخاطر قطع شدن لینک (از Hub میان راه)، یا برفرض خاموش شدن یک سوییچ دیگر BPDU ی دریافت نمیکند و به Max-age میرسد، پورت اول آن Root Port میشود.

برای دیگر سگمنت، پورت دوم Cat-C امکان Designated Port شدن را دارد تمام این اتفاقات در کمتر از یک دقیقه رخ میدهند و در مجموع 50 ثانیه (20 Max Age + 15 Listening + 15 Learning) به حالت Forwarding درمیاید. زمان Convergence در STP، معمولا 30 تا 50 ثانیه است.

Convergence در شبکه حالتی است که شبکه پس از ناپایداری به حالت پایدار میرسد.

تغییر تایمر تنها روی Root Bridge امکان پذیر است و در صورتیکه روی یک سوییچ دیگر این زمان بندی را تغییر دهید از آن صرف نظر خواهد شد. اما باید توجه داشت که روی Backup Root Bridge نیز زمانبدنی لحاظ گردد. هر چند که تغییر مقادیر پیش فرض پیشنهاد نمیشود مگر با علم به جزییات STP.

پیغام های STP

دو گونه Message در STP بین سوئیچ ها رد و بدل میشود:

- Configuration BPDU
- Topology Change Notification (TCN) BPDU

- Configuration BPDU از Root Bridge به بقیه ارسال میشود.
- TCN از سوئیچ ها به سمت Root Bridge موقع تغییرات ارسال میشود.
- در یک شبکه سالم و صحیح، اکثر پیغام های STP از نوع Configuration BPDU هستند.

در جدول زیر فیلد های مختلف یک Configuration BPDU ذکر شده است.

Field	بایت	توضیحات
Protocol ID	2	همیشه صفر است.
Version	1	همیشه صفر است.
Type	1	برای Configuration BPDU برابر صفر و برای TCN برابر یک است.
Flags	1	LSB = Topology Change (TC) flag MSB = Topology Change Acknowledgment (TCA) flag
Root BID	8	Bridge ID مربوط به Root Bridge فعلی.
Root Path Cost	4	Root Path Cost تا Root Bridge را مشخص میکند.
Sender BID	8	فرستنده BPDU را با BID آن مشخص میکند.
Port ID	2	مشخص کننده پورت ارسال کننده BPDU
Message Age	2	زمان تولد Message در Root Bridge
Max Age	2	طول عمر صحت یک Configuration BPDU
Hello Time	2	زمانبندی بین ارسال BPDU Configuration ها
Forward Delay	2	زمانبندی Listening و Learning

یک TCN تنها شامل سه فیلد اول است.

Cisco in Persian

تغییر Topology در STP

در موقع بروز هرگونه اشکال و یا تغییر وضعیت پورت یا اتصال یک سگمنت، STP باید با خبر شده و دست به کار شود و محاسبات دوباره صورت گیرد. سوئیچ تغییر را از Root Port به سمت Root Bridge ارسال میکند. هر سوئیچ میانه، تغییر را دریافت کرده و به فرستنده ACK میفرستد.

وقتی تغییری در وضعیت پورت سوئیچ داده شود بصورت TCN آنرا به Root Port ارسال میکند تا Root متوجه تغییر شود.

STP برای از کار انداختن Loop تنها با State و وضعیت پورتها بازی میکند. پورت نهایتا در وضعیت blocking یا forwarding کار میکند.

هرگاه تغییری در STP رخ دهد (در واقع State و وضعیت پورتی تغییر کرده است) سوئیچ تغییر را توسط ارسال TCN یا Topology Change Notification اطلاع میدهد. این TCN حاوی جزئیات تغییر نیست، بلکه خبر وقوع تغییر است. TCN از Root Port سوئیچ مربوطه به سمت Root Bridge ارسال میگردد.

Upstream Switch (سوییچ بالاتر) به محض دریافت TCN، ACK یا رسید به فرستنده ارسال میکند و سوییچ زیری تا زمانی که ACK را دریافت نکند به ارسال TCN هر دو ثانیه (Hello Time) ادامه میدهد. این Hello Time تنظیم شده روی خود سوییچ است و میتواند از Root Bridge Hello Time متفاوت باشد.

Cisco in Persian

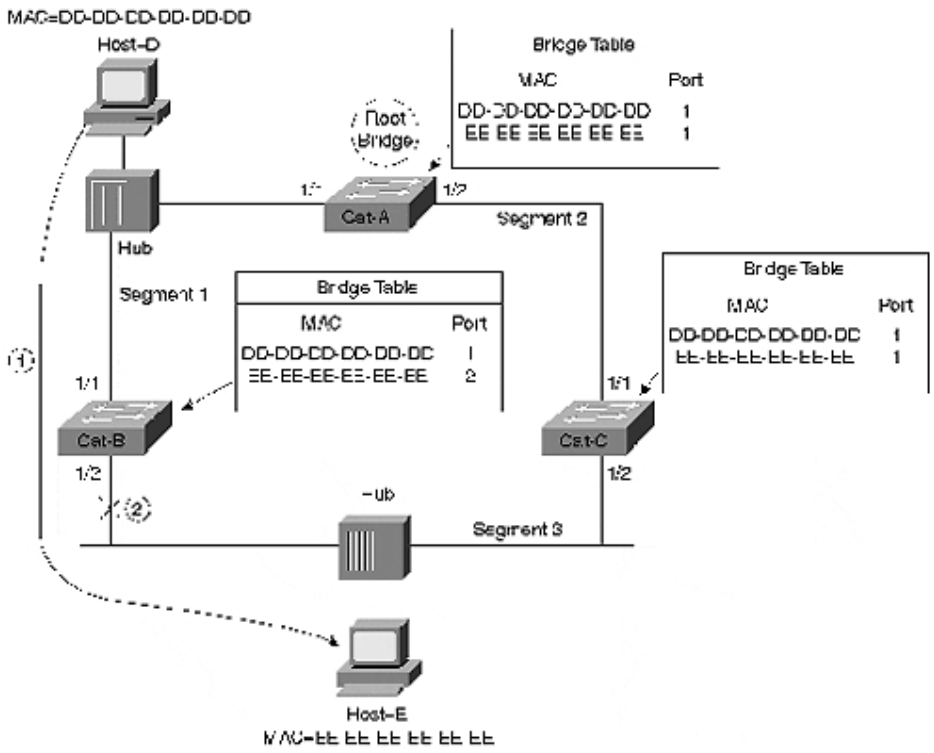
سوئیچ بالایی (upstream switch) در ارسال Configuration BPDU، قسمت Topology Change Acknowledge را علامت میزند و رسیدن TCN را به فرستنده اطلاع میدهد تا دیگر TCN نفرستد.

حال سوئیچ بالایی TCN ایجاد کرده و به سوئیچ بالاتر از خود آنرا ارسال میکند، با این تفاوت که به Root Bridge یک گام نزدیکتر شده ایم. مراحل رسید و ارسال ACK مطابق بالا تکرار میگردد.

Root Bridge به محض دریافت TCN، Topology Change ACK را ارسال میکند ضمن اینکه هنگام ارسال BPDU Configuration بعدی، Topology Change Flag را علامت میزند. این کار را بمدت زمان واحدی برابر با Forward Delay + Max Age یعنی 15+20=35 ثانیه در BPDU Configuration های ارسالی تکرار میکند.

فایده ارسال این علامت این است که:

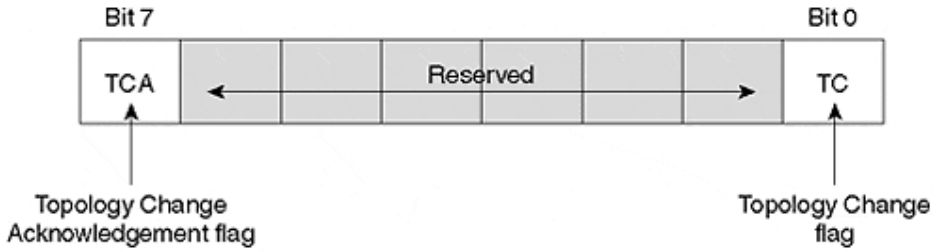
سوئیچ ها به محض دریافت Topology Change Flag در Configuration BPDU، زمان طول عمر جدول Bridging را به میزان Forward Delay کاهش میدهند. یعنی از 300 ثانیه به 15 ثانیه کاسته و اگر در این زمان فریمی از MAC Address خاصی دریافت نشد از جدول حذف میشود. این کار موجب بهبود Convergence Time در شبکه میگردد.



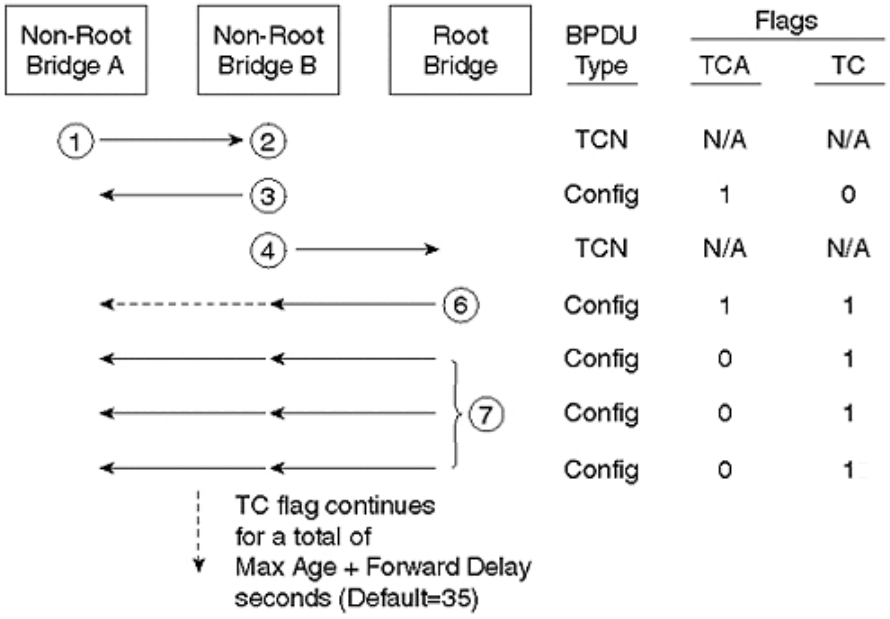
Cisco in Persian

طبق شکل بالا، تغییر وضعیت یک پورت و ارسال TCN باید موجب تغییر Bridging Table نیز شود.

Flag های Configuration BPDU در زیر آمده است:



مراحل ارسال TCN و خبر وقوع تغییر در شبکه، در شکل زیر نشان داده شده است:



Cisco in Persian

STP به ازای هر VLAN

هر آنچه که تا بحال در رابطه با STP گفتیم، مستقلا برای هر VLAN اجرا میگردد. پیاده سازی STP برای VLAN توسط سیسکو و IEEE متفاوت است.

CST (Common Spanning Tree) در استاندارد 802.1Q تعریف شده و به پیاده سازی تنها یک STP به ازای تمامی VLAN ها اشاره میکند. CST BPDU ها بصورت untagged و در Native VLAN مبادله میگردند. مزیت CST سادگی و مصرف پایین منابع (Resources) در شبکه است.

اما CST با اشکالاتی نیز همراه است:

در حالتی خاص ممکن است مسیر Trunk بین راه، عبور فریم های یک VLAN را پشتیبانی نکند در نتیجه ارتباط درون یک VLAN بدرستی برقرار نگردد.

در مقابل سیسکو به ازای هر VLAN یک STP جداگانه اجرا میکند یعنی سوییچی میتواند در VLAN 2 نقش Root Bridge را ایفا کند در حالیکه در VLAN 3 سوییچ دیگر این کار را انجام دهد. برای هر VLAN یک درخت جداگانه شکل میگیرد که مختص به همان شبکه لایه دو است.

Cisco in Persian

Per-VLAN Spanning Tree یا **PVST** به ازای هر VLAN یک STP اجرا میکند و توسط Cisco ارائه شده است. PVST برای کار به Trunk های ISL سیسکو نیاز دارد و این خود مشکل است و موجب میشود با CST اصلا سازگار نباشد.

Per-VLAN Spanning Tree Plus یا **PVST+** توسط سیسکو برای سازگاری بین متد های مختلف نظیر PVST, CST و PVST+ ارائه شد و روی ترانک های 802.1Q نیز کار میکند.

تعیین Root Bridge

قرارگیری و موقعیت منطقی Root Bridge در شبکه حائز اهمیت است زیرا کل گراف شبکه، از Root به سمت شاخه ها شکل میگیرد. پس با استفاده از تغییر اولویت، میتوان شانس سوئیچ خاص را برای Root شدن را بالا برد.

با هم عرض بودن سوئیچ ها و داشتن اولویت مساوی, ممکن است یک سوئیچ لایه Access در مرکز شبکه به جای Core ترافیک زیادی را متحمل شود!

برای حل این مساله دو کار باید انجام دهیم:

1- سوئیچی را بعنوان Root Bridge تعیین و تنظیم کنیم.
2- سوئیچ دیگری را بعنوان Secondary Root Bridge برای موقعی که Root Bridge دچار مشکل شد، انتخاب و تنظیم نماییم.

یکی از فاکتورهای مهم در قرارگیری Root Bridge مرکزیت آن در شبکه است. از دیگر فاکتورهای مهم نزدیکی سوئیچ به Server Farm است.

Cisco in Persian

به دو صورت انتخاب Root Bridge قابل تنظیم است:

1- پائین آوردن اولویت با دستور:

```
spanning-tree vlan vlan priority priority
```

2- استفاده از یک ماکرو بصورت دستور زیر:

```
spanning-tree vlan vlan root {primary|secondary} [diameter diameter]
```

این ماکرو چندین کار انجام میدهد، ابتدا زمانبندی STP را بصورت پیش فرض در میارد.

سپس با اولویت بازی میکند تا سوییچ, Root شود. البته هیچ تضمینی برای Root شدن سوئیچ در صورتیکه باقی سوییچ ها با اولویت های مختلف و غیر Default تنظیم شده باشند وجود ندارد و روش اول پیشنهاد میگردد.

در زمان اجرای ماکرو, سوئیچ اولویت Root Bridge فعلی را میسنجد و اگر اولویت آن زیر 24576 بود اولویت خود را 4096 ست میکند در غیر این صورت اولویت خود را 24576 تنظیم خواهد کرد.

برای Secondary Root Bridge، اولویت 28672 تنظیم میگردد با این فرض که اولویت همه سوییچ ها در حالت default, 32768 است. هیچ مکانیزمی برای تعیین اولویت Secondary Root Bridge وجود ندارد و این اعداد تنها قراردادی بوده و توسط سیسکو انتخاب شده اند.

اصلی ترین کار در تنظیم STP همین انتخاب Root Bridge است، باقی کارها به گونه ای خودکار توسط STA یا Spanning Tree Algorithm صورت میگیرد. اما گاهی اوقات ممکن است، بین انتخاب مسیرهای چندگانه هم عرض با Cost برابر بخواهیم فرقی قائل شویم. باید توجه داشت که STP در زمان تصمیم گیری از شروط چهارگانه ای که بعنوان معیار های STP شرح دادیم به ترتیب زیر استفاده میکند:

1- پائین ترین BID
2- پائین ترین Root Path Cost
3- پائین ترین Sender BID
4- پائین ترین Port ID

انعطاف پذیری در تنظیمات STP

زمانی که سوئیچ BPDU را دریافت میکند، ارزش پورت خود را به ارزش دریافتی روی BPDU اضافه کرده و Root Path Cost را محاسبه میکند. این Cost به ازای پورت و VLAN قابل تنظیم است:

```
spanning-tree [vlan vlan] cost cost
```

Cisco in Persian

در صورتیکه سوییچ با دو پورت به یک سگمنت وصل شود و پورت ها ارزش یکسان داشته باشند از Port ID برای انتخاب یکی و Block کردن دیگری کمک میگیرد. Port ID تنها یک عدد اختصاص داده شده به پورت است. Port ID قابل تنظیم بوده و این پارامتر از 16 بیت تشکیل میشود:

8 بیت Port Priority یعنی اولویت + 8 بیت Port Number که با توجه به قسمت اولویت (Priority) میتوان Port ID را تغییر داد. عدد پورت یا Port Number از یک تا 255 بوده و غیر قابل تغییر است اما عدد اولویت بصورت پیش فرض 128 است.

```
spanning-tree [vlan vlan] priority priority
```

برای تغییر پارامتر های زمان بندی STP از دستور های زیر استفاده میکنیم:

```
spanning-tree [vlan vlan] hello-time seconds (1~10 – default=2)
spanning-tree [vlan vlan] forward-time seconds (4~30 – default=15)
spanning-tree [vlan vlan] max-age seconds (6~40 – default=20)
```

بهبود Convergence در STP

همانطور که اشاره شد Convergence در شبکه حالتی است که شبکه پس از ناپایداری به حالت پایدار میرسد و زمان Convergence هرچه کمتر باشد بهتر است. از طرق مختلف زمان Convergence در سطوح access، uplink و حتی backbone را میتوان بهبود بخشید:

- **PortFast** برای اتصال سریع کامپیوترها به سوئیچ های access.
- **UplinkFast** برای بکار افتادن لینک redundant در سوئیچ access به مرکز شبکه.
- **BackboneFast** برای بهبود convergence در backbone بعد از وقوع تغییرات.

PortFast

برای بالا آمدن یک پورت و رسیدن به وضعیت forwarding، دو زمان learning و listening هر کدام 15 ثانیه و در مجموع 30 ثانیه وقت صرف میشود. اگر از تکنولوژی های دیگر نظیر PAgP نیز روی لینک استفاده کرده باشیم (20 ثانیه تاخیر مربوط به ether-channel) این زمان به 50 ثانیه طول خواهد کشید که برای یک پورت متصل به یک Computer بسیار طولانی است.

Cisco in Persian

با تنظیم PortFast روی پورت هایی که به Workstation ها متصلند و نیازی به احتیاط STP نبوده و پورت مراحل Listening و Learning را طی نمیکندو زمان اتصال پورت مطلوب میشود.

با وجود PortFast، پورت مستقیما به حالت forwarding بالا می آید.

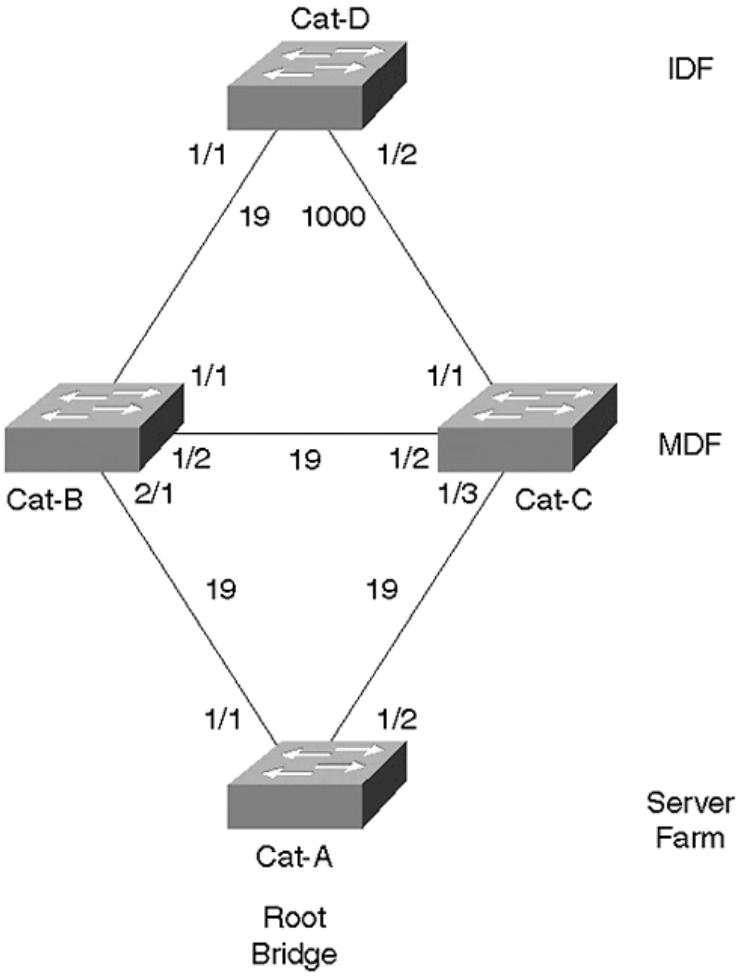
از مزایای PortFast این است که به ازای تغییر وضعیت پورت، TCN ایجاد نمیگردد.

```
Switch(config-if)# spanning-tree portfast
```

UplinkFast

در حالتی که سوئیچ access به دو distribution switch متصل باشد، پورت redundant با صرف 50 ثانیه تاخیر از قطع شدن پورت اصلی شروع به کار میکند. با UplinkFast این زمان به حداقل ممکن رسیده و uplink طی چند ثانیه up میشود. روی سوئیچ های access تنها باید UplinkFast را فعال کنیم.

UplinkFast پورت کاندید Root Port را به حالت Backup Root Port درآورده تا در زمان نیاز بسرعت لینک برقرار گردد.



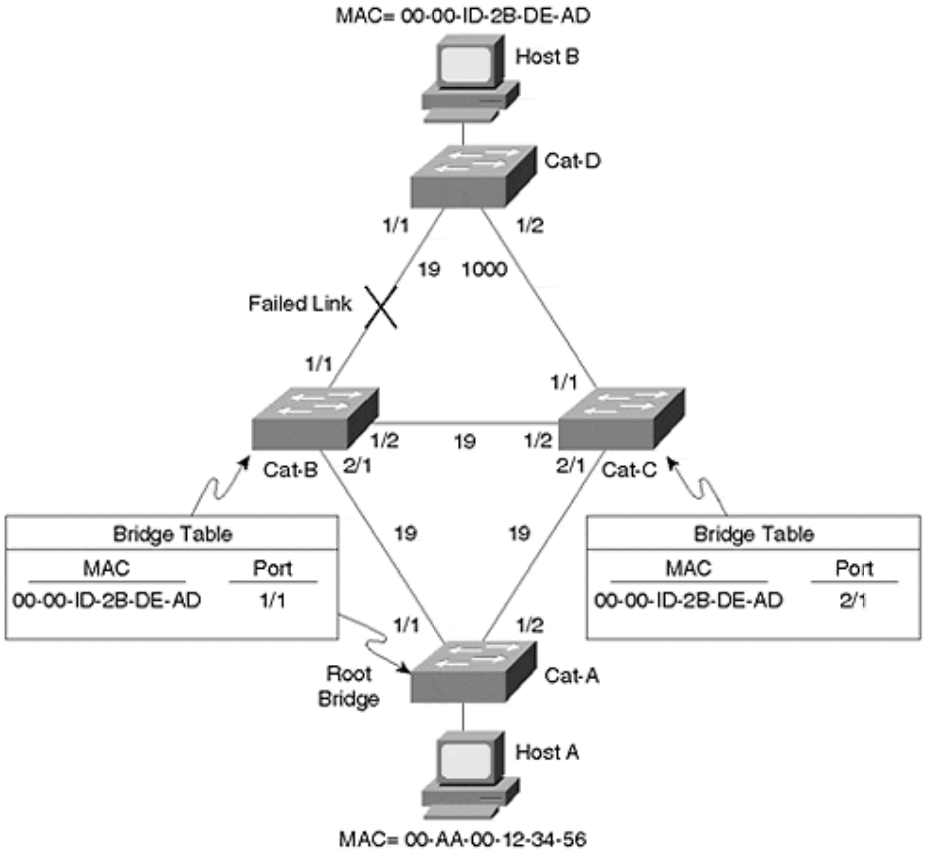
در شکل بالا، شبکه از بخش های IDF یا Intermediate Distribution Frame که همان access است و MDF یا Main Distribution Frame که همان Core/Distribution است تشکیل شده است.

Cat-D برای اتصال از پورت 1/1 استفاده میکند اما پورت دوم بحالت آماده باش برای uplink شدن سریع است.

دستور زیر UplinkFast را در سوئیچ فعال میکند:

```
Switch(config)# spanning-tree uplinkfast [max-update-rate pps]
```

بمحض وارد شدن این دستور UplinkFast در تمام VLAN ها و کل سوئیچ اعمال میگردد. تمایل سوئیچ برای Root Bridge شدن کم شده و Priority به 49152 میرسد. برای اینکه سوئیچ Transit بقیه سوئیچ ها نشود نیز به هزینه پورت ها 3000 واحد افزوده میشود. این دستور روی Root Bridge مجاز نیست.



و اما قسمت آخر دستور فوق بعنوان Max-update-rate چه فایده ای دارد؟

سوئیچ در هنگام تغییر لینک، Bridge Table خود و پورت خروجی را بسادگی تغییر میدهد اما باید به سوئیچ بالاتر نیز این تغییر و جابجائی MAC Address ها را از لینکی به لینک دیگر گزارش دهد تا عملیات سرعت گیرد.

این عملیات با ارسال یک سری Multicast Frame به آدرس 0100.0ccd.cdcd که آدرسی ساختگی است، از آدرس فرستنده های داخل CAM (Content Addressable Memory) صورت گرفته و سرعت ارسال این فریم ها در واحد ثانیه با **max-update-rate** تنظیم میشود که بصورت پیش فرض 150pps است.

BackboneFast

BackboneFast برای بهبود زمان Max Age در نظر گرفته شده و نهایتا زمان Convergence را از 50 ثانیه به 30 ثانیه کاهش میدهد. البته BackboneFast روی تمامی سوئیچ ها باید فعال شود که بصورت پیش فرض غیر فعال است.

BackboneFast قطع شدن یک سوئیچ را از روی لینک دیگر سوییچ ها تشخیص میدهد و به آن تشخیص indirect link-failure میگویند. این تشخیص از دریافت Inferior BPDU قابل استنباط است.

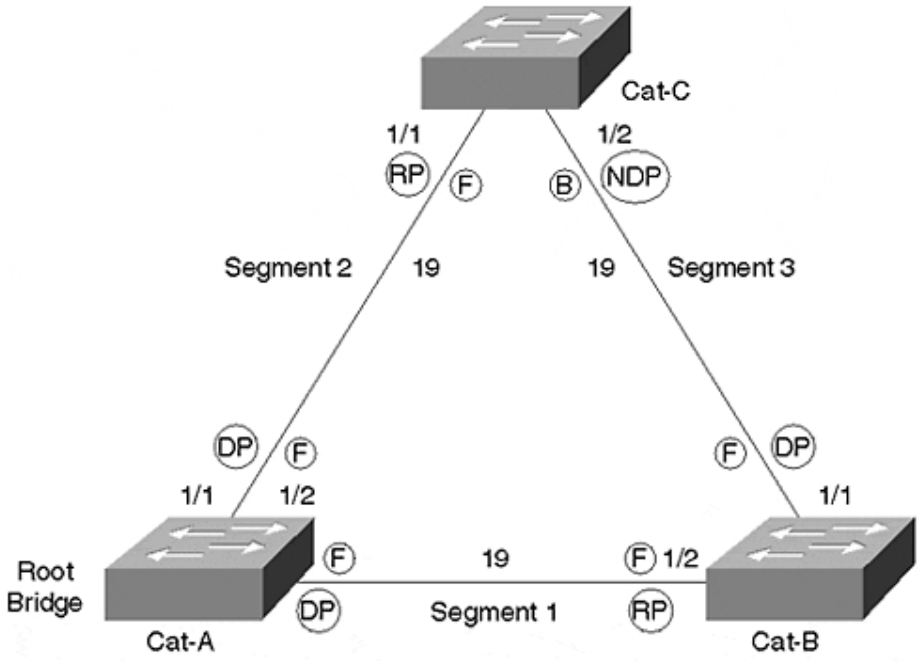
Inferior BPDU ها، از طرف سوئیچی ارسال میگردند که ارتباط خود را با Root Bridge از دست داده و خود را Root اعلام میکند.

در حالت عادی، برای اینکه سوئیچ به Inferior BPDU ها گوش دهد و یا خود را Root اعلام کند، باید Max Age سپری شود اما BackboneFast این زمان را از بین برده و موجب تسریع عملیات میگردد.

سوئیچ RLQ یا Root Link Query میفرستد تا Root Bridge را پیدا کند، اگر جواب را روی Root Port خود دریافت کند مسیر سالم است و اگر روی پورتهای block شده دریافت کرد، یک Root Port دیگر باید انتخاب کند و Max Age را صفر میکند.

Cisco in Persian

سوئیچ بالاتر که RLQ را دریافت میکند در صورتیکه Root Bridge باشد یا Root Bridge را گم کرده باشد، RLQ Reply میفرستد، در غیر اینصورت RLQ را به سوئیچ های بالاتر خود فرستاده تا Reply برای فرستنده ایجاد و فرستاده شود.



در شکل بالا Cat-B وقتی Segment 1 را قطع شده میبیند خود را Root Bridge اعلام میکند. اما Cat-C با ارسال RLQ ارتباط خود با Root را مطمئن شده و به Inferior BPDU ها عمل نمیکند و در نهایت Root Port در Cat-B تغییر میکند.

Cisco in Persian

مقابله با BPDU های مزاحم

- **Inferior BPDU**: BPDU ی که از سمت یک سوئیچ که Root خود را گم کرده ارسال میشود. در این پیام سوییچ خود را Root اعلام میکند.
- **Superior BPDU**: سوییچی که این پیام را ارسال میکند خود را Root اعلام میکند و از BID بهتری نسبت به Root فعلی برخوردار است پس مستحق Root شدن است اما اگر این BDPU از سوی سوییچ یک مشتری یا کاربر باشد میتواند مخرب باشد زیرا توپولوژی شبکه ما را به هم میریزد و Root شبکه ما شده و ترافیک قابل توجه ای را ترانزیت خواهد کرد.

برای مقابله با انواع مخرب BPDU، از دو راه حل میتوان استفاده کرد:

- **Root Guard**: وقتی در پورتی تنظیم شود، در صورت دریافت Superior BPDU، پورت به وضعیت root-inconsistent در می آید و در واقع پورت بلوکه میشود. این وضعیت زمانی از بین میرود و به حالت نرمال در می آید که دیگر از Superior BPDU ها خبری نباشد.

  در واقع با این تکنیک محدوده ای برای Root شدن تعیین میکنیم (Root Guard بصورت پیش فرض غیرفعال است.)

```
Switch(config-if)# spanning-tree guard root
```

Cisco in Persian

- **BPDU Guard**: اگر هرگونه BPDU از پورتی که BPDU Guard دارد دریافت شود پورت فورا در وضعیت errdisable قرار میگیرد. این خصوصیت بصورت پیش فرض غیرفعال است و پیشنهاد میشود روی پورت هایی که PortFast دارند، فعال گردد زیرا از این پورت ها انتظار دریافت BPDU را نداریم.

```
Switch(config-if)# spanning-tree bpduguard enable
```

در صورتیکه BPDU Guard و PortFast روی پورت متصل به هاب (Hub) تنظیم شوند از Loop شدن جلوگیری نخواهد شد چون Hub با ما Spanning Tree صحبت نمیکند و خود STP ندارد.

باید توجه داشت که PortFast به منزله غیرفعال کردن STP روی پورت نیست.

مقابله با گم شدن BPDU

سیسکو برای BPDU های مزاحم راه حل Root Guard و BPDU Guard را ارائه داده است، اما در حالتی که اصلا BPDU دریافت نشود چه راهکاری در شبکه موجود است؟

BPDU Skew Detection

تاخیر دریافت BPDU ها را محاسبه میکند و گزارش میدهد. زمان تاخیر بعنوان skew time اندازه گیری میشود و از طریق Syslog گزارش میشود.

Loop Guard

برای جلوگیری از Loop شدن یک سگمنت در حالیکه Root را پیدا نمیکند ارائه شده است. این خصوصیت بصورت پیش فرض غیرفعال است و روی پورت هایی که Designated port نیستند عموما استفاده میشود مثل Uplink ها تا اگر Root بخاطر تاخیر یا هر علتی پیدا نشد و زمان به Max-Age رسید، از Forwarding شدن آن پورت جلوگیری شده و شبکه بخاطر تاخیر در STP، Loop نشود.

توسط دستور زیر روی پورت فعال میشود:

Cisco in Persian

```
Switch(config-if)# spanning-tree guard loop
```

Unidirectional Link Detection

در شرایط خاصی نظیر استفاده از فیبر نوری ممکن است ارتباط یک طرفه شود اما دو طرف از آن مطلع نگردند بطور مثال درست کارنکردن یک GBIC ممکن است لینک را connected نشان دهد اما در طرفی فریم ها دریافت یا ارسال نشوند. این لینک unidirectional یا یک طرفه برای STP خطرناک است زیرا BPDU ها در سمت دیگر دریافت نمیشوند.

در چنین حالتی باید از UDLD یا Unidirectional Link Detection که در STP درنظر گرفته شده استفاده کنیم. بدین صورت که در بازه های زمانی مشخص (15 ثانیه) سوییچ یک فریم مخصوص ارسال میکند و از طرف مقابل انتظار دارد تا همان فریم را همراه با مشخصه پورت آن، باز پس (Echo) بفرستد.

UDLD در دو سر لینک روی دو سوئیچ باید تنظیم گردد.

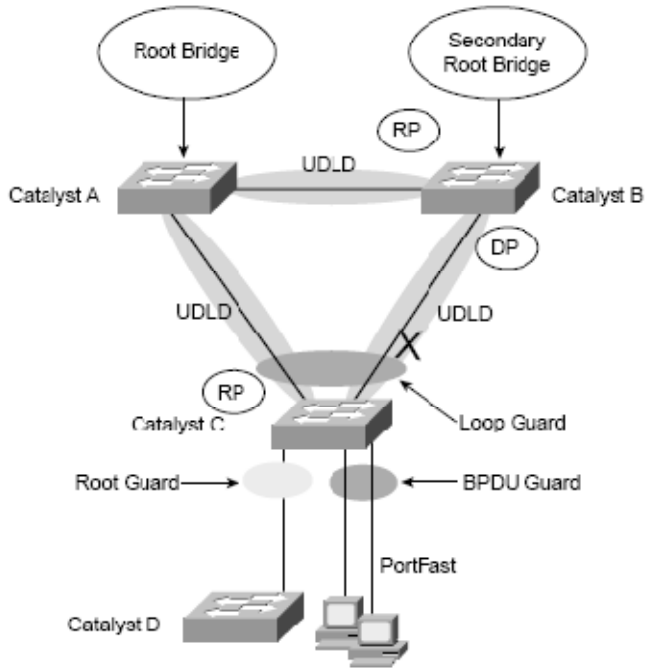
UDLD در دو مد Normal و Aggressive کار میکند. در مد Normal اگر لینکی یک طرفه شناخته شد، مجاز است بکار خود ادامه دهد اما پیغام اخطار به Syslog ارسال خواهد شد. در مد Aggressive سوئیچ سعی میکند ارتباط را برقرار کند و زمان ارسال پیام UDLD را به 8 ثانیه کاهش میدهد. اگر هیچ کدام

echo نشد، پورت یکطرفه مجاز به کار نخواهد بود و errdisable میشود. UDLD بصورت پیش فرض غیرفعال است. برای فعال کردن آن در سوئیچ از دستور زیر استفاده میکنیم:

```
Switch(config)# udld {aggressive | enable | message time seconds}
```

با دستور فوق سوئیچ UDLD را تنها برای پورت های فیبر فعال میکند و در صورت تمایل روی هر پورتی نظیر پورت های Copper قابل فعال شدن است:

```
Switch(config-if)# udld {aggressive | disable | enable}
```



Root guard: Apply to ports where root is never expected.

BPDU guard: Apply to all user ports where PortFast is enabled.

Loop guard: Apply to nondesignated ports; but okay to apply to all ports.

UDLD: Apply to all fiber optic links between switches (must be enabled on both ends).

Permissible Combinatons on a Switch port:

Loop guard and UDLD
Root guard and UDLD

Not permissible on a switch port:

Root guard and Loop guard
Root guard and BPDU guard

RTSP

IEEE 802.1w

تغییرات توپولوژی در STP عموما تا 30 ثانیه طول کشیده تا شبکه خود را درمان کند و به حالت Loop-free درآید. استاندارد IEEE 802.1w برای بهبود بخشیدن به 802.1D تدوین شد. زمانیکه سیسکو امکاناتی نظیر Portfast و Uplinkfast را ارائه کرد استاندارد 802.1D نیز نیاز داشت تا خود را بروز کند.

Rapid Spanning Tree Protocol برای ارائه روشی جهت ارتباط سوئیچ ها در شبکه Loop-free و در عین حال سرعت بخشیدن به زمان Convergence، تدوین شد. این پروتکل که براساس همان STP است در یک VLAN یا چند VLAN همراه با MST یا MSTP یا Multiple Spanning Tree (استاندارد IEEE 802.1s) بکار گرفته میشود.

در 802.1D نقش یک پورت یکی از سه حالت: Root Port, Designated Port و یا Blocked (در حالتی که پورت Root یا Designated نباشد) بود:

- Root Port
- Designated Port
- Blocking Port

Cisco in Persian

در 802.1D وضعیت هر پورت نیز یکی از حالات زیر است:

- Disabled
- Blocking
- Listening
- Learning
- Forwarding

انتخاب Root Bridge در پروتکل جدیدتر 802.1w درست همانند 802.1D است و براساس کوچکترین BID است اما نقش پورت در RSTP بصورت مقابل است:

- Root Port
- Designated Port
- Alternate Port
- Backup Port

- Root Port: بهترین Root Path Cost را دارد. Root Bridge هیچ Root Port ی ندارد.
- Designated Port: پورتی در سگمنت که بهترین Root Path Cost را داراست.
- Alternate Port: پورتی جانشین برای Root Port که مسیر خوبی به Root دارد اما از مرغوبیت پائینتری نسبت به Root Port برخوردار است.
- Backup Port: پورتی که مرغوبیتی کمتر نسبت به Designated Port دارد و در صورت قطع ارتباط سگمنت, "ممکن است" یک مسیر به Root داشته باشد.

وضعیت پورت ها در RSTP بر اساس رفتار با فریم های ورودی تعریف میگردد و طبق نقش پورت، یکی از سه حالت زیر خواهد بود:

- Discarding
- Learning
- Forwarding

- Discarding وضعیتی است که فریم های ورودی دور انداخته میشوند و هیچ آدرس MAC ی روی پورت یادگیری نمیشود. این مرحله تلفیقی از مراحل disabled, blocking, listening در 802.1D است.
- Learning: آدرس های MAC را یاد میگیرد اما کماکان انتقالی صورت نمیگیرد.
- Forwarding: فریم ها بر اساس Bridging Table منتقل میشوند.

وضعیت (Status) پورت در RSTP و STP در جدول زیر مقایسه شده است:

STP	RSTP	توضیح
Disabled	Discarding	پورت فریم ها را دور می اندازد.
Blocking		
Listening		
Learning	Learning	پورت به یادگیری MAC ها میپردازد.
Forwarding	Forwarding	پورت به وضعیت عادی بر میگردد.

Cisco in Persian

نقش هر پورت در STP و RSTP با یکدیگر در جدول زیر مقایسه شده است:

STP	RSTP
Root Port	Root Port
Designated Port	Designated Port
Blocking Port	Alternate Port
	Backup Port

وضعیت BPDU ها در RSTP

فرم BPDU ها در RSTP سازگار و مطابق با BPDU های STP است. شماره نسخه BPDU از صفر به 2 تبدیل شده و از بیتهای رزرو در RSTP برای مصارف مختلف استفاده میشود (STP BPDU Version=0)

همسایگان در RSTP بر سر وضعیت پورت ها با هم توافق (negotiate) میکنند. زمان Hello همان 2 ثانیه است و ارسال و دریافت BPDU بین همسایگان حتی اگر Configuration BPDU از Root نرسد, ادامه خواهد داشت و این کار برای نگهداری توپولوژی شبکه الزامی است. نقش و وضعیت پورتهای سوئیچ توسط flag های مختلف به همسایگان گزارش میشود.

وقتی سه BPDU متوالی, دریافت نشود (6 ثانیه) سوئیچ همسایه را مرده تشخیص میدهد. (بر عکس STP که Max-Age آن 20 ثانیه است)

RSTP در کنار STP میتواند همزیستی داشته باشد و اگر روی پورتی STP دریافت کند، آن پورت STP گونه رفتار خواهد کرد. این عملکرد به ازای هر پورت است و پس از زمان خاصی پورت دوباره سعی میکندبه حالت RSTP در آمده و با طرف مقابل صحبت کند و اگر مجددا STP BPDU گرفت عملیات تکرار میشود.

Cisco in Persian

انواع پورت در RSTP

Edge Port: پورتی که به یک Host متصل شده و اگر BPDU روی پورت دریافت شود، به حالت non-edge تبدیل میشود. این خاصیت در STP نیز مثل RSTP بوسیله PortFast روی پورت فعال میشد.

```
Switch(config-if)# spanning-tree portfast
```

Root Port: پورتی که بهترین مسیر به Root را داراست و در صورتیکه پورت های دیگری نیز در سوئیچ فعال باشند (مسیرهای مختلف) تحت عنوان Alternate Port در صورت لزوم، به Forwarding در آمده و کار خود را شروع میکنند.

Point-to-Point Port: این گونه پورت ها لینک بین سوئیچ ها بوده و در واقع Shared نیستند. در این حالت یکی از سوئیچ ها درخواست Designated شدن را میدهد و دیگری براساس مقادیر خودش پیشنهاد را قبول یا رد میکند. این لینک Point-to-Point به ارتباطات Full-Duplex اشاره میکند و شامل Half-Duplex ها نمیشود.

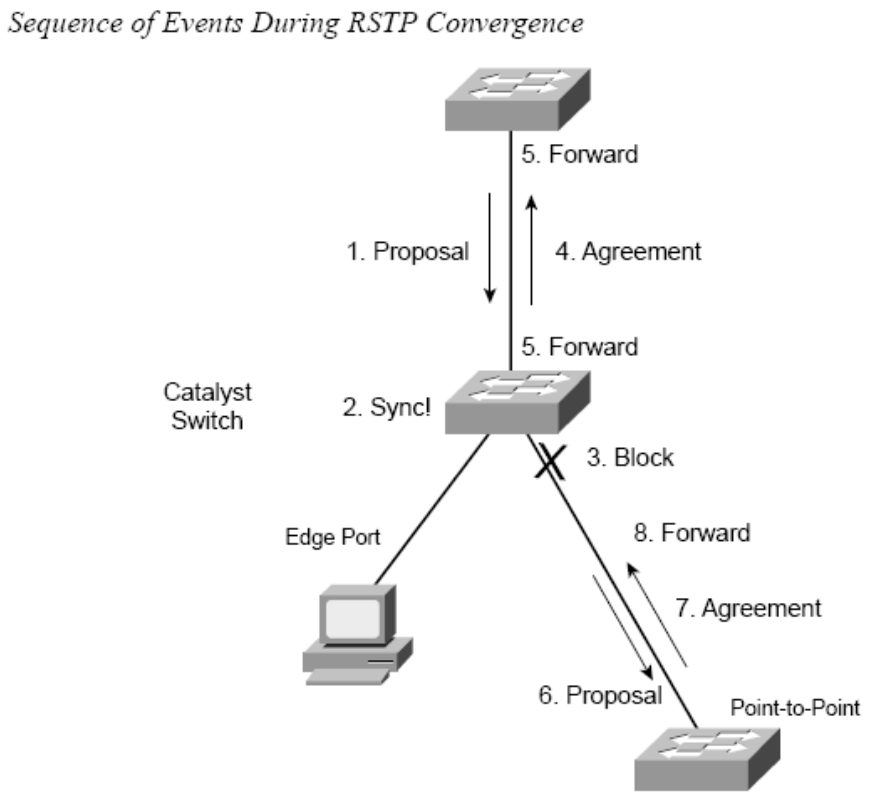
برای تنظیم یک پورت به حالت Point-to-Point از دستور زیر استفاده میشود:

```
Switch(config-if)# spanning-tree link-type point-to-point
```

Cisco in Persian

RSTP Synchronization

طبق مراحل زیر Convergence در RSTP شکل میگیرد:

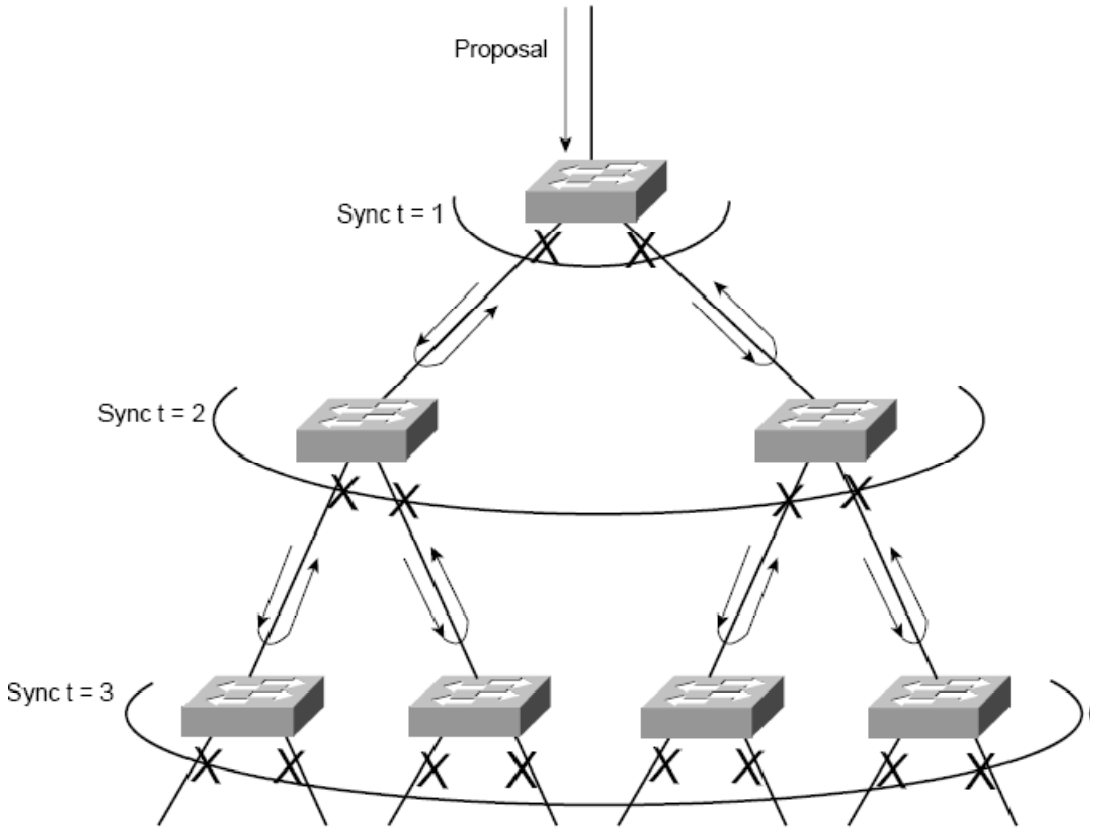


1- بین سوییچ ها برای Designated شدن پورت هایشان روی پورت های غیر edge مذاکره صورت میگیرد. (Proposal-Agreement Handshake) بدین صورت که فرستنده ی Superior BPDU، Designated شده و پورت مقابل Root Port میشود.
2- سوئیچ خود را با شبکه و توپولوژی Sync میکند.
3- بقیه پورت هایی که edge نیستند به حالت blocking درمی آیند.
4- Agreement و توافق به ارسال کننده Proposal، فرستاده میشود (بصورت Configuration BPDU و به فرستنده میگوید که در حال Sync است.)
5- هر دو پورت؛ Root Port و Designated Port به حالت Forward درمیایند.
6- به باقی پورت ها که در مد discarding هستند Proposal ارسال میشود. (غیر از edge ها)
7- منتظر پیغام Agreement از همسایه میماند تا دریافت شود.
8- پورت به حالت Forwarding تغییر وضعیت میدهد.

تغییر توپولوژی در RSTP

وقتی در شبکه تغییری رخ میدهد، RSTP به ازای تغییر در وضعیت یک non-edge port، Topology Change یا TC ارسال میکند. (یعنی برای edge port ها این اتفاق نمی افتد) بیت TC در BPDU ست شده و از non-edge designated port ها ارسال میگردد.

همسایگانی که TC را گرفتند، همه MAC های درون Bridge Table خود را flush میکنند، غیر از MAC هایی که به فرستنده TC ارجاع داده خواهد شد.



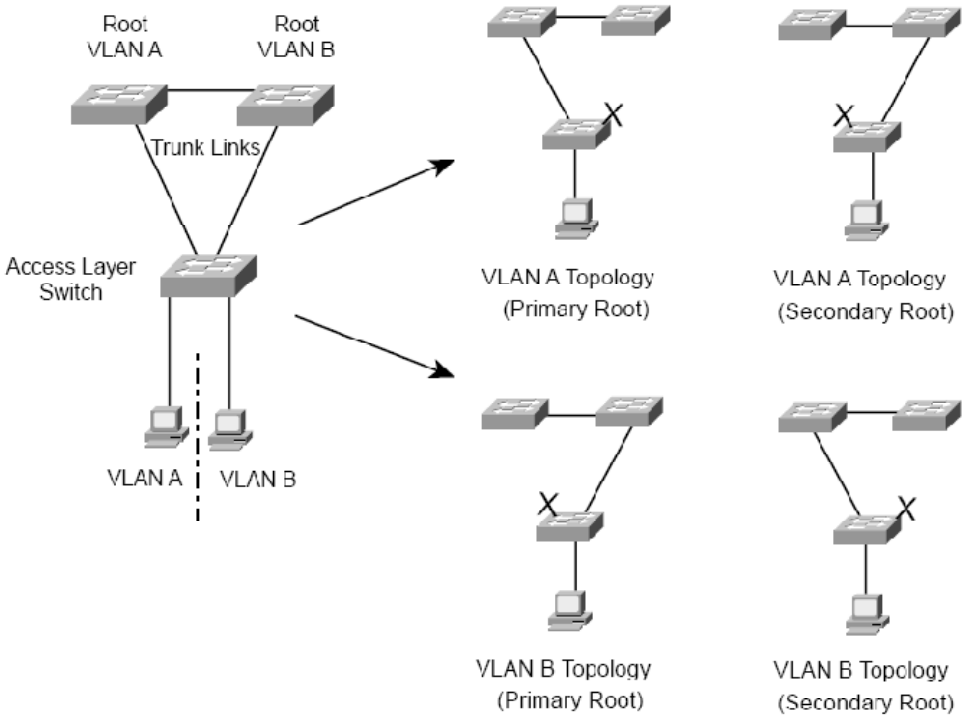
در حالت default، سیسکو از STP همراه با PVST+ استفاده می کند.

Multiple Spanning Tree

IEEE 802.1s

در STP بصورت سنتی, از CST (Common Spanning Tree) یا در سیسکو از PVST+ استفاده میشود، CST به ازای همه VLAN ها یک مسیر Loop-free توسط یک پروسه STP ایجاد میکند و این مسیر نمیتواند مسیر بهینه برای همه VLAN ها باشد لذا توپولوژی شبکه به ازای هر VLAN کاراترین نخواهد بود.

PVST+ نیز همیشه بهترین راه حل نیست. از آنجا که به ازای هر VLAN، STP مستقلی ایجاد میکند و این کار باعث صرف زیاد منابع میشود در شبکه های بزرگ و پر VLAN میشود یعنی اشغال CPU، حافظه و همچنین پهنای باند. اگر یکصدVLAN داشته باشیم، یکصد STP با گراف مختص به خود و Root Bridge جداگانه تشکیل میگردند در حالیکه بعید نیست تعداد توپولوژی های ممکن شبکه تنها دو حالت باشد!



به تعداد VLAN ها، STP داریم اما تعداد توپولوژی ها (گراف ها) محدود است.

با MST، چند STP Instance (گراف) ایجاد میشود و VLAN ها به Instance ها Map میشوند.

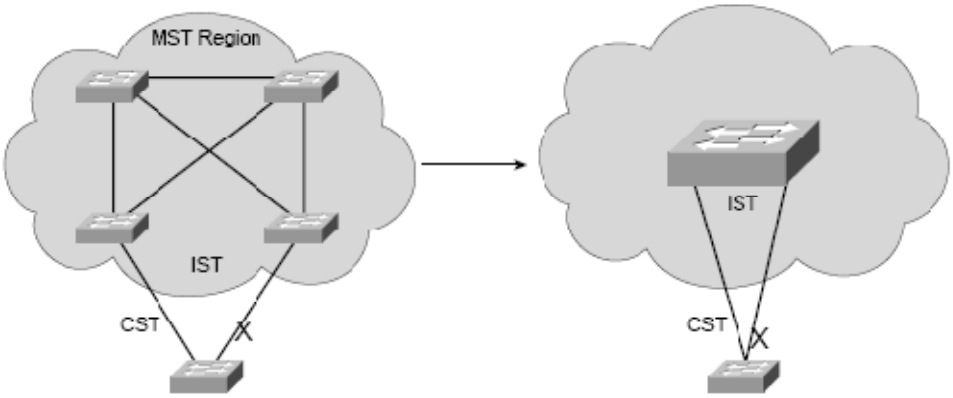
MST Region

گرچه MST بسیار متفاوت با CST و +PVST است اما توانایی سازگاری و برقراری ارتباط با آنها را دارد. تمام سوئیچ های یک ناحیه باید MST را اجرا کنند و پارامتر های زیر بین سوئیچ ها یکسان باشد تا در یک ناحیه محسوب شوند:

- MST name (32 Character)
- MST Revision number (2 bytes)
- MST instance-to-VLAN mapping table (4096 entries)

هر Instance-to-VLAN mapping در همه سوئیچ ها باید تنظیم گردد و این جدول توسط BPDU ها منتقل نمیشود. در عوض Digest (کد محاسبه شده از محتویات جدول) بین سوئیچ ها رد و بدل میشود تا از یکسان بودن جداول و Mapping مطمئن شوند.

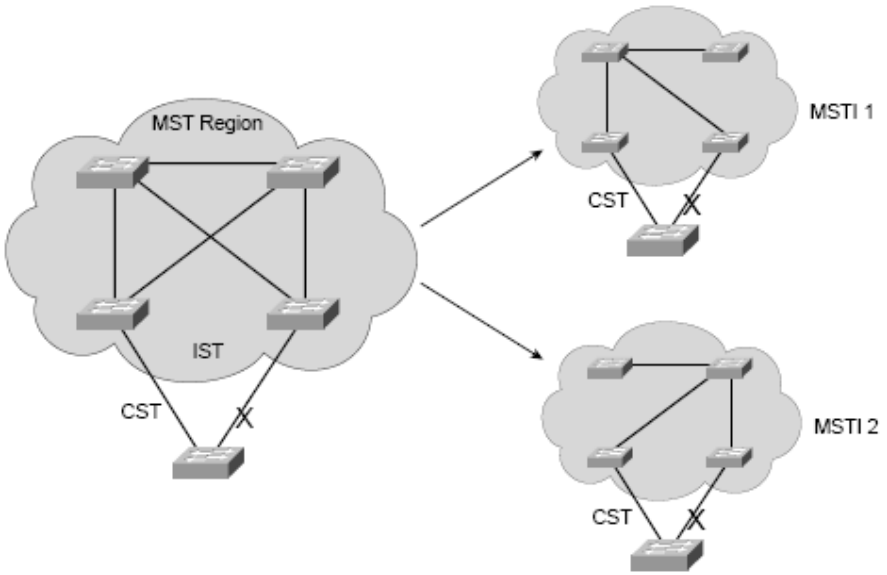
برای اینکه از یک ناحیه اطلاعات STP به خارج ارسال شود از CST استفاده میشود و اطلاعات کلیه VLAN ها بصورت untagged روی Native-VLAN به بیرون ارسال و از بیرون دریافت میگردد. CST شبکه داخلی را بعنوان یک Instance داخلی یا IST یا Internal Spanning Tree میبیند.



در شکل بالا، سوئیچ پائینی از CST استفاده میکند پس مجموعه بالا را یک Instance یا نمونه میبیند و پورت دوم خود را برای جلوگیری از Loop به حالت blocking در می آورد.

در MST، Instance های واقعی در کنار IST فعال میشوند و VLAN ها به MST Instance یا MSTI، Map میشود. سیسکو تا 16 عدد MSTI را پشتیبانی میکند که MSTI 0 به IST تعلق دارد و MSTI ها از یک تا 15 را در اختیار دارند.

در شکل زیر سه نوع نمونه یا Instance داریم, MSTI 1, MSTI 2 و IST.



BPDU های کل MST تنها از طریق IST (MSTI 0) ردوبدل میشود. اطلاعات هر MSTI با MST BPDU بوسیله M-record قابل شناسایی است و به ازای کل 16 نمونه ممکن تنها به یک BPDU نیاز داریم.

Cisco in Persian

تنها IST BPDU ها در بیرون MST Region های مختلف رد و بدل میشوند.

اگر MST Region اطلاعاتی دریافت کند که مربوط به چند VLAN باشد (با دریافت BPDU های متفاوت) به این نتیجه میرسد که به PVST+ متصل شده است. برای ارسال اطلاعات خود به PVST+، IST از هر BPDU یک نسخه به هر VLAN در PVST+ Trunk ارسال میکند.

بصورت پیش فرض تمام VLAN ها به IST یا Instance 0، Map شده اند.

تنظیمات MST

برای بکارگیری MST در شبکه، MST باید روی هر سوییچ تنظیم شود.

```
Switch(config)# spanning-tree mode mst
Switch(config)# spanning-tree mst configuration
Switch(config-mst)# name name
Switch(config-mst)# revision version-number
Switch(config-mst)# instance instance-id vlan vlan-list
```

از دستور های بالا میتوان نتیجه گرفت که برای تنظیم MST باید به مد Configuration مخصوص MST رفت و اسم و revision number را در همه سوئیچ ها یکسان تنظیم کرد. هنگامیکه تغییری در تنظیمات دادیم باید یک واحد به revision number افزود و این کار نیز در تک تک سوئیچ های MST Region باید انجام شود.

برای دیدن تغییرات در حال اعمال از دستور زیر استفاده میکنیم:

```
Switch(config-mst)# show pending
```

برای خروج از تنظیمات MST از exit استفاده میکنیم.

```
Switch(config-mst)# exit
```

Cisco in Persian

بعد از تنظیم MST، سوئیچ PVST+ را غیرفعال کرده و RSTP را فعال میکند.

جدول تنظیمات MST در زیر نشان داده شده است: (Timer ها به IST مرتبط بوده و در کل MST یکسان خواهند بود)

دستور	تنظیم مربوط به
spanning-tree mst *instance-id* **root {primary | secondary} [diameter** *diameter***]**	Root Bridge
spanning-tree mst *instance-id* **priority** *bridge-priority*	Bridge Priority
spanning-tree mst *instance-id* **cost** *cost*	Port Cost
spanning-tree mst *instance-id* **port-priority** *port-priority*	Port Priority
spanning-tree mst hello-time *seconds* **spanning-tree mst forward-time** *seconds* **spanning-tree mst max-age** *seconds*	STP Timers