

Cisco APIC M1/M2/L1/L2 から M3/L3 クラスタへの移行

バージョン 1.1

このドキュメントの目的

このドキュメントでは、旧世代の APIC サーバーを L3/M3 モデルにインサービス交換する方法について詳しく説明します。cisco.com¹で発表されたように、APIC L1/M1 と APIC L2/M2 の両方のサーバーが販売終了日とサポート終了日を迎えました。このドキュメントの作成時点で推奨されている Cisco APIC サーバーは、APIC L3/M3 です。

ソフトウェア リリース要件

APIC L3/M3 には、Cisco APIC ソフトウェア 4.0(1) 以降のリリース、または 3.2(8) 以降の 3.2 リリースが必要です。このドキュメントでは、Cisco APIC 4.0(3d) リリースを例として使用します。クラスタを構成する APIC サーバーは、すべて同じソフトウェア リリースを実行する必要があります。1 つのクラスタ内で異なるソフトウェア リリースを使用することはできません。これに従わないと、クラスタが収束しなくなります。このルールには 1 つの例外があります。ソフトウェア アップグレードプロセス中は、クラスタ内のソフトウェア リリースに一時的な相違が生じます。それで、既存の APIC M1/L1 または APIC M2/L2 を APIC M3 / L3 サーバーに置き換える前に、実行中のクラスタをサポートされているリリースにする必要があります。

APIC M3/L3 サーバーで現在実行しているリリース バージョンを確認するには、次の手順を実行します。

ステップ 1. APIC M3/L3 の電源をオンにし、現在実行しているリリースを確認します。すぐに使用できる APIC M3/L3 では、4.0 リリース トレインの最新リリースであるリリース 4.0(3d) が実行されます。

APIC M3/L3 が 4.0(3d) を実行していない場合は、サーバーをサポートされているリリースにアップグレードする必要があります。

ステップ 2. この手順は、Cisco Integrated Management Controller (CIMC) の仮想コンソールまたは Serial over LAN (SoL) 接続を使用して、cisco.com から取得した ISO イメージから直接起動できます。²

ステップ 3. 続行する前に、既存のクラスタを同じリリースにする必要があります。

ハードウェアの互換性

APIC M1/L1、APIC M2/L2、および APIC M3/L3 は組み合わせて使用できます。ソフトウェア リリース要件に記載されている最小ソフトウェア リリース以外の制限はありません。

表 1. Table Caption

	APIC-M1/L1	APIC-M2/L2	APIC-M3/L3
APIC-M1/L1	X	X	X
APIC-M2/L2	X	X	X
APIC-M3/L3	X	X	X

¹ <https://www.cisco.com/c/en/us/products/cloud-systems-management/application-policy-infrastructure-controller-apic/eos-eol-notice-listing.html>

² https://community.cisco.com/legacyfs/online/attachments/discussion/technote-enablesolonapic_v1.pdf

クラスタが異なるハードウェアモデルを混在させると、そのパフォーマンスは最小のものに合わせられます。たとえば、APIC-M3 クラスタはエッジポート数を 1200 にまで拡張できるのに対し、APIC-M2 クラスタは 1000 までです。³

インサービス APIC サーバーの交換

ここでは、データプレーンやコントロールプレーンに影響を与えずに、既存のすべての M1/M2/L1/L2 サーバーを、稼働中の M3/L3 サーバーモデルに置き換える方法について説明します。この手順は Cisco によりフルサポートされています。この手順は 3 ノード APIC クラスタに焦点を当てていますが、プロセスは大規模なクラスタでも同様です。

注意

作業を開始する前に、次を確認してください。

- 一度に複数の APIC をデコミッションしないこと
- クラスタが完全な適合状態になるまで待つから、次の交換に進んでください。
- デコミッションされた APIC の電源をオンのままにしないでください。
- ソフトウェアリリースが一致しない新しい APIC を導入しない

ステップ 1. 既存のクラスタが完全に一致していることを検証します。

この手順を実行する前に、既存のクラスタが完全に適合していることを確認してください。完全に適合していない APIC クラスタをアップグレードまたは変更してはなりません。既存のクラスタが完全に適合していることを確認するには、次の手順を実行します。

- メインページで、[コントローラ (Controllers)] を選択します。
- [コントローラ (Controllers)] を展開し、任意の APIC を選択します。
- APIC を展開し、[クラスタのノード表示 (cluster as seen by node)] を選択します。

図 1

ID	Name	IP	Admin State	Operational State
1	apic1	10.0.0.1	In Service	Available
2	apic2	10.0.0.2	In Service	Available
3	apic3	10.0.0.3	In Service	Available

- すべてのノードの動作状態を確認します。使用可能な状態であり、正常性が完全な適合状態であると報告される必要があります。

³ https://www.cisco.com/c/ja_jp/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/datasheet-c78-739715.html

- e. 図 2、図 3、および図 4では、最初のクラスタに 1 つの APIC M1 と 2 つの APIC M2 が含まれており、クラスタは 4.0(3d) を実行しています。

図 2

```
Hardware Usage
Allocated Memory (KB): 23981460
Free Memory (KB): 41585496
APIC Model: APIC-SERVER-M1
APIC Serial Number: FCH1829V2YW
CPU Architecture: x86_64
Cores: 6
CPU Model: Intel(R) Xeon(R) CPU E5-2620 v2 @ 2.10GHz
Speed (GHz): 2.1
Vendor: GenuineIntel
Locator LED: Off
Locator LED Color: No color
```

図 3

```
Hardware Usage
Allocated Memory (KB): 36573024
Free Memory (KB): 28856612
APIC Model: APIC-SERVER-M2
APIC Serial Number: FCH2010V1AH
CPU Architecture: x86_64
Cores: 6
CPU Model: Intel(R) Xeon(R) CPU E5-2609 v3 @ 1.90GHz
Speed (GHz): 1.2
Vendor: GenuineIntel
Locator LED: Off
Locator LED Color: No color
```

図 4

```
Hardware Usage
Allocated Memory (KB): 35073976
Free Memory (KB): 30357560
APIC Model: APIC-SERVER-M2
APIC Serial Number: FCH2036V3K1
CPU Architecture: x86_64
Cores: 6
CPU Model: Intel(R) Xeon(R) CPU E5-2609 v3 @ 1.90GHz
Speed (GHz): 1.895
Vendor: GenuineIntel
Locator LED: Off
Locator LED Color: No color
```

ステップ 2. 交換用の APIC M3/L3 サーバーをケーブル接続します。

このシナリオでは、2 台の APIC サーバー、つまり M1 と 1 台の M2 のみを交換します。このプロセスは、3 台、5 台、または 7 台のサーバーすべてを交換する場合でも同じです。データセンターに交換用サーバーを物理的に設置し、他のサーバーと同様に既存の ACI ファブリックにケーブル接続します。必要に応じて、CIMC NIC レベルで LLDP が無効になっていることを確認します。⁴アウトオブバンド (OOB) 管理接続をケーブル接

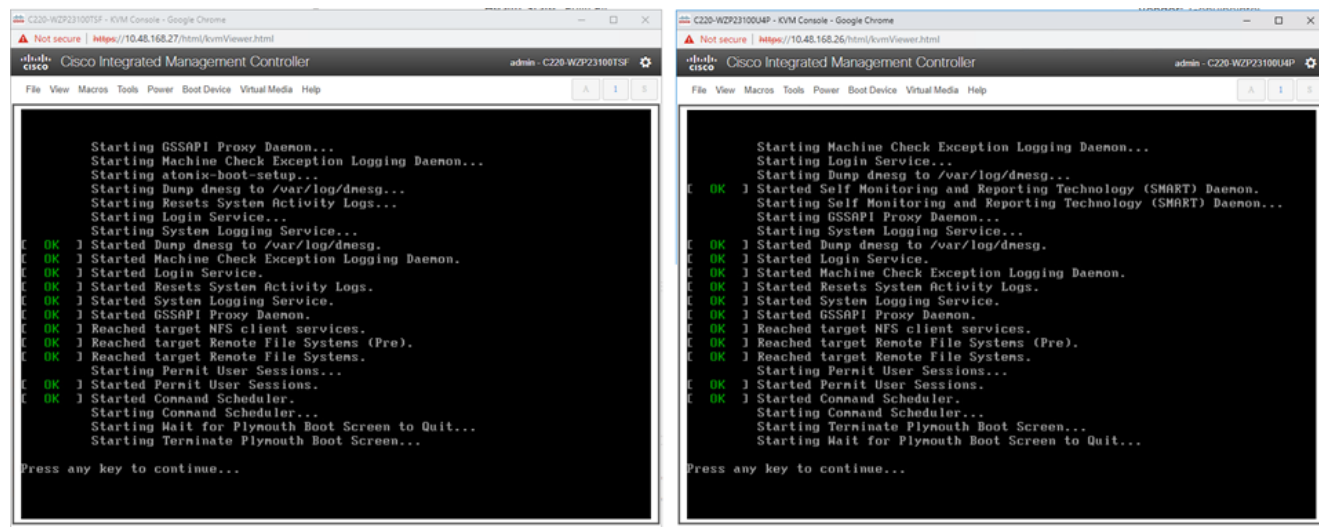
⁴ <https://community.cisco.com/legacyfs/online/attachments/document/files/apic-vic-lldp-fn.pdf>

続します。各 APIC は交換するサーバーの IP を単に引き継ぐため、交換用の APIC サーバーに新しい IP アドレスを設定する必要はありません。

ステップ 3. 交換用 APIC サーバー M3/L3 の電源を入れます。

すべての APIC M3/L3 サーバーの電源を投入し、仮想キーボード、ビデオ、マウス セッション、Serial over LAN (SoL)、または物理 VGA 接続を起動して、ブートプロセスをモニタできるようにします。数分後、いずれかのキーを押して続行するように求められます。しかし、まだキーは押さないでください。この時点では、APIC M3/L3 サーバーはその段階のままにします。図 5 を参照してください。

図 5

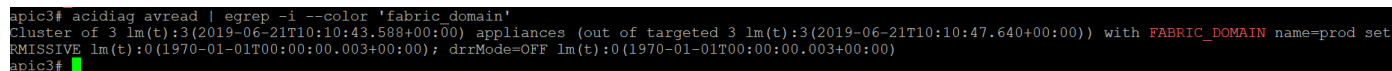


ステップ 4. 既存のファブリックの名前とインフラ VLAN を記録します。

現在のファブリックと、使用中のインフラ VLAN の正確な名前を確認しておく必要があります。図 6 は、ファブリックの名前を確認する方法を示しています。ファブリックの名前を確認するには、既存のいずれかの APIC で次のコマンドを入力します。

```
apic3# acidiag avread | egrep -i -color 'fabric_domin'
```

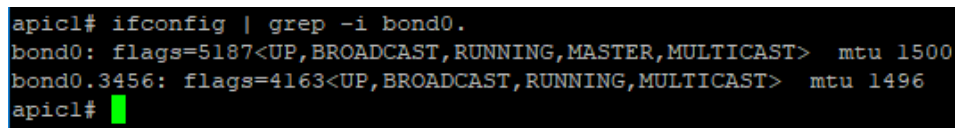
図 6



現在のドメインの名前を記録します。図の例では [prod] です。

- 使用しているインフラ VLAN がもうない場合は、`ifconfig | grep -i bond0` コマンドを使用して検索します。この例では、Infra VLAN は 3456 です。

図 7



- また、最初にファブリックを起動したときに使用した TEP プールとマルチキャスト プールも必要です。TEP プールを取得するには、次のコマンドを入力します。

```
apic1# acidiag avread | grep -o -P 'tep address=.{0,18}'
```

- マルチキャスト IP プールを決定するには、次のコマンドを入力します。

```
apic1# moquery -c fvBD | grep -E "name|bcastP|dn" | grep -B 2 "infra"
```

Error! Reference source not found. に例を示します。

ステップ 5. APIC 3、5、または 7 をデコミッションします。

[クラスタのノード表示 (cluster as seen by node)] ビュー (図 1) の APIC 1 または 2 から最後の APIC を右クリックして、デコミッション (decommission) を選択します (**Error! Reference source not found.**)。

図 8

from APIC-2, you can decommission APIC-1 or APIC-3

ID	Name	IP
1	apic1	10.0.0.1
2	apic2	10.0.0.2
3	apic3	10.0.0.3

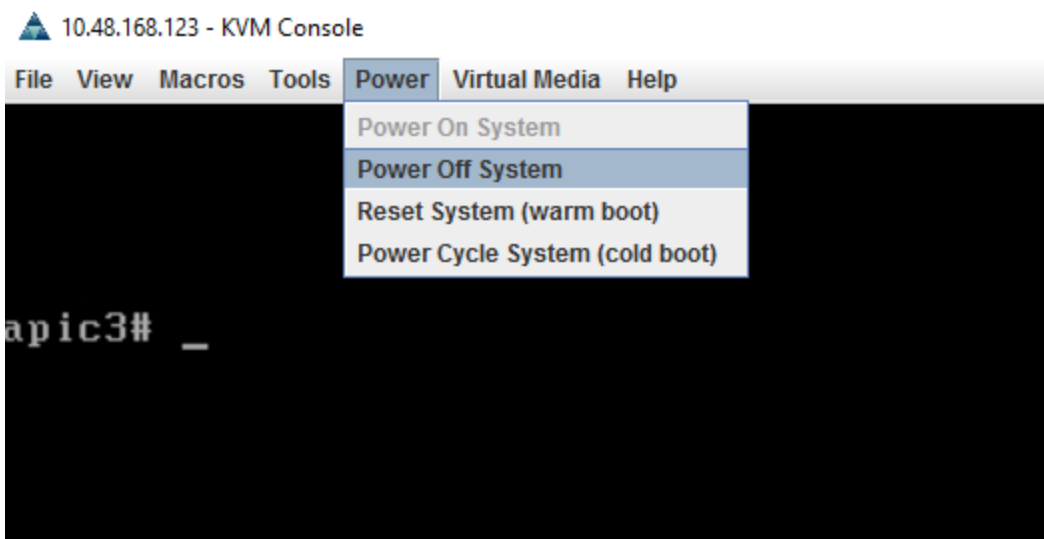
約 5 分待つてから、APIC の CIMC にログインするか、背面に物理キーボードとモニタを接続して、APIC サーバーの使用停止後に電源オフシーケンスを開始できるようにします。ステータスがインサービスからアウトオブサービスに変わります。

図 9

ID	Name	IP	Admin State	Operational State
1	apic1	10.0.0.1	In Service	Available
2	apic2	10.0.0.2	In Service	Available
3	apic3	10.0.0.3	Out of Service	Available

古い APIC がアウトオブサービスの場合は、電源をオフにします。

図 10



しばらくすると、図 11 に示すように、この APIC のステータスが未登録に変わります。

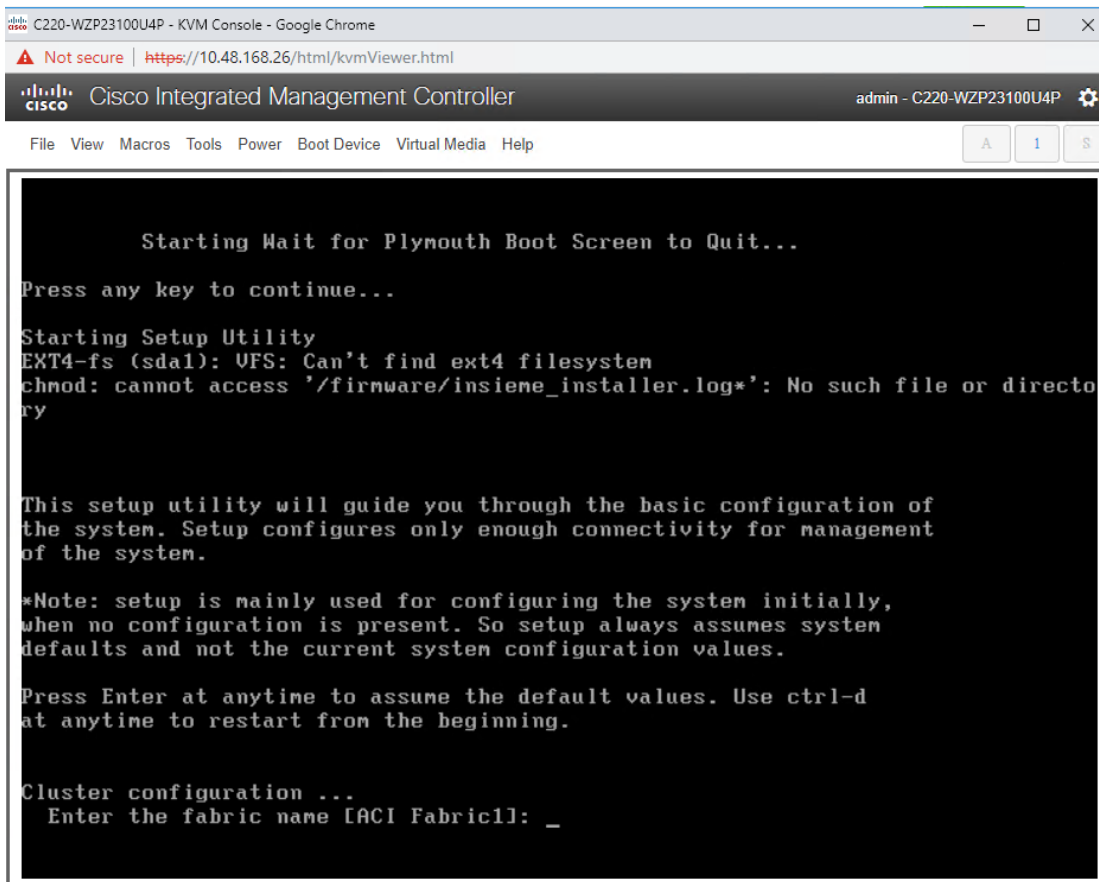
図 11

Active Controllers					
ID	Name	IP	Admin State	Operational State	Health State
1	apic1	10.0.0.1	In Service	Available	Fully Fit
2	apic2	10.0.0.2	In Service	Available	Fully Fit
3	apic3	0.0.0.0	Out of Service	Unregistered	Unknown

ステップ 6. 交換用 APIC を持ち込みます。

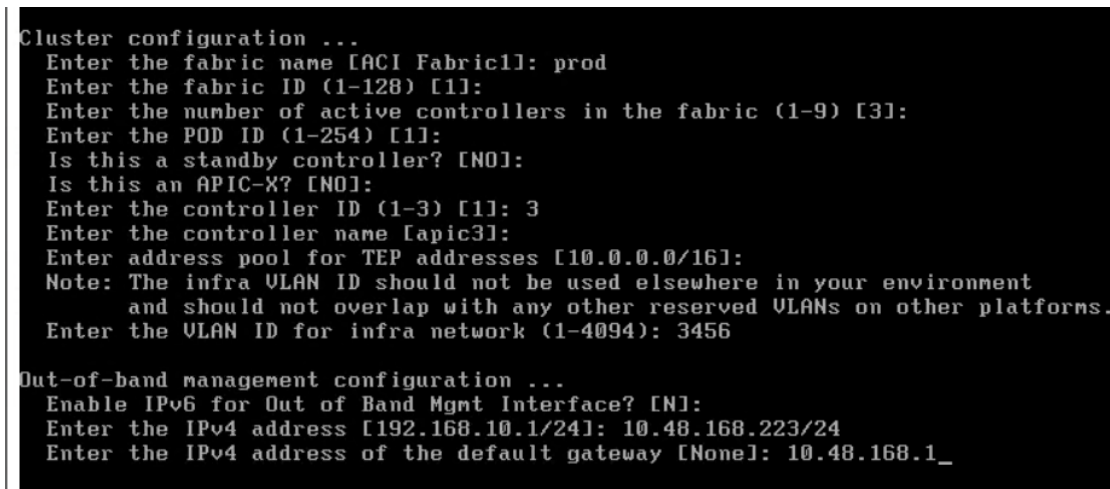
「press any key to continue」プロンプトが表示されたままになっていた、新しい APIC M3/L3 のうちの 1 台を選択します。ここまで押さないで来ましたが、この時点でいずれかのキーを押して続行することができます。この APIC を設定するように求められます。ステップ 4 で収集したファブリックの名前を入力し、交換する APIC とまったく同じ設定パラメータを適用します。たとえば、名前が「apic3」で、IP が 10.48.168.223 だった番号 3 の APIC を置き換える場合は、新しい APIC に図 12 と図 13 に示すこれらのパラメータを使用します。

図 12



ステップ 4 で取得した値を使用して、新しい APIC の設定を開始します。

図 13



すべてのパラメータを入力すると、それらを変更するかどうかを尋ねられます。間違いを修正したい場合を除き、「N」と応答します。

ステップ 7. 新しい APIC をクラスタ メンバーシップに登録します。

約 7~10 分経過すると、新しいサーバーは、次のように UI の [クラスタのノード表示 (cluster as seen by node)] タブに未登録状態で表示されます。サーバーを右クリックしてコミッションします。新しいサーバーの正常性がすべてのサーバーに完全に適合するまで待つてから、続行します。これには通常 5 分かかります。

図 14

Active Controllers

ID	Name	IP	Admin State	Operational State
1	apic1	10.0.0.1	In Service	Available
2	apic2	10.0.0.2	In Service	Available
3	apic3	0.0.0.0	Out of Service	Unregistered

Commission
Decommission
Replace

ステップ 8. クラスタ メンバーシップを検証します。

5 分程度経過すると、動作状態と正常性ステータスの遷移が観察されます。新しいサーバーは、完全に収束する前に、最初にデータ層が部分的に分岐した状態になります。

図 15

Active Controllers

ID	Name	IP	Admin State	Operational State	Health State	Failover Status	Serial Number
1	apic1	10.0.0.1	In Service	Available	Fully Fit	idle	FCH2036V3K1
2	apic2	10.0.0.2	In Service	Available	Fully Fit	idle	FCH2010V1AH
3	apic3	10.0.0.3	In Service	Available	Data Layer Partially Diverged	idle	WZP23100U4P

その後すぐに、新しいサーバーのデータベースがクラスタの他のメンバーと完全に同期します。このことは、正常性が完全に適合したとして反映されます。

図 16

Active Controllers

ID	Name	IP	Admin State	Operational State	Health State	Failover Status	Serial Number
1	apic1	10.0.0.1	In Service	Available	Fully Fit	idle	FCH2036V3K1
2	apic2	10.0.0.2	In Service	Available	Fully Fit	idle	FCH2010V1AH
3	apic3	10.0.0.3	In Service	Available	Fully Fit	idle	WZP23100U4P

新しいサーバーのプロパティを拡大すると、新しいシリアル番号を持つ M3/L3 であることがわかります。

図 17

Hardware Usage

Allocated Memory (KB): 24588768

Free Memory (KB): 72759092

APIC Model: APIC-SERVER-M3

APIC Serial Number: WZP23100U4P

CPU Architecture: x86_64

Cores: 8

CPU Model: Intel(R) Xeon(R) Bronze 3106 CPU @ 1.70GHz

Speed (GHz): 0.8

Vendor: GenuineIntel

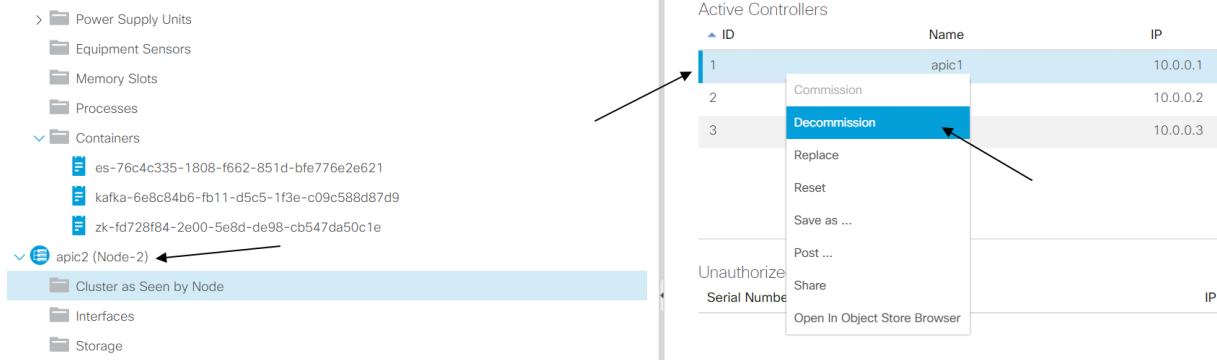
Locator LED: Off

Locator LED Color: No color

ステップ 9. 別のサーバーをデコミッションします。

別のサーバーをデコミッションするには、ステップ 5-8 を繰り返します。コントローラをデコミッションする場合、その操作は別のサーバーから実行する必要があることに注意してください。たとえば、APIC-1 にログインしている状態で、APIC-1 をデコミッションしないでください。APIC-2 にログインし、APIC-2 の [クラスタのノード表示 (cluster as seen by node)] ビューに移動し、APIC-1 をデコミッションします。これを次に示します。

図 18



交換を行う前に、デコミッションにしたサーバーの電源をオフにすることを忘れないでください。

ステップ 10. クラスタ全体を確認します。

サーバーをデコミッションして電源をオフにした後、**M3** を起動、設定、コミッションし、必要な回数だけルーティングを行います。クラスタ全体が完全に適合していることを検証します。

図 19

ID	Name	IP	Admin State	Operational State	Health State	Fallover Status	Serial Number
1	apic1	10.0.0.1	In Service	Available	Fully Fit	idle	WZP23100TSP
2	apic2	10.0.0.2	In Service	Available	Fully Fit	idle	FCH2010V1AH
3	apic3	10.0.0.3	In Service	Available	Fully Fit	idle	WZP23100U4P

APIC-1 の交換対象も **M3** モデルです。

図 20

Hardware Usage

Allocated Memory (KB): 31056080

Free Memory (KB): 66291780

APIC Model: APIC-SERVER-M3

APIC Serial Number: WZP23100TSP

CPU Architecture: x86_64

Cores: 8

CPU Model: Intel(R) Xeon(R) Bronze 3106 CPU @ 1.70GHz

Speed (GHz): 1.44

Vendor: GenuineIntel

Locator LED: Off

Locator LED Color: No color

この時点で、新しいハードウェアを備えた完全に機能する **APIC** クラスタが完全に動作します。

通常のクラスタで置き換えられる、スタンバイ **APIC** サーバーのデコミッション

クラスタに古いスタンバイ **APIC** サーバーが含まれている場合は、同じプロセスを適用します。既存のクラスタを、サポートされているリリースに移行すると、スタンバイ **APIC** サーバーも自動的にアップグレードされません。

スタンバイ **APIC** サーバーをデコミッションするには、次の手順を実行します。

ステップ 1. 新しい **M3** または **L3** モデルが、他のクラスタ メンバーと同じソフトウェア リリースを実行していることを確認します。

ステップ 2. 通常のクラスタメンバーと交換するために、スタンバイ APIC をデコミッションします。電源を切り、コントローラが登録解除されるまで待ちます。

ステップ 3. 新しい M3 または L3 サーバーを持ち込み、セットアップ時にサーバーがスタンバイ APIC として指定します。"Is this a standby controller?[NO]"（「これはスタンバイ コントローラですか? [いいえ]」）というプロンプトが表示されるので、次のように入力します。

Is this a standby controller? [NO]: (ここまでがプロンプト。デフォルトは NO) YES (「YES」と入力)

新しいクラスタのトラブルシューティング

新しいクラスタメンバーはクラスタに参加できないのは、ほとんどの場合、インフラ VLAN、TEP プール、ファブリック名、およびマルチキャストプールの設定パラメータが正しくないか、ケーブル接続が正しくないためです。これらを再確認する必要があります。新しいコントローラが完全に収束するまでに少し時間がかかることに注意してください。少なくとも 10 分待ちます。Rescue-User アカウントを使用すれば、準備ができていないクラスタメンバーにいつでもログインできます。クラスタが検出モードの場合、パスワードは不要です。パスワードが必要な場合は、admin パスワードを使用します。

ステップ 1. ファブリックへの物理インターフェイスを確認します。

ファブリックへのインターフェイスが稼働していることを確認します。「cat /proc/net/bonding/bond0」コマンドを入力します。少なくとも 1 つのインターフェイスがアップしている必要があります。これは、クラスタメンバーシップを確立するために必要かつ十分な条件です。ただし、単一のインターフェイスだけが稼働している場合は、APIC で重大な、またはクリティカルな障害が発生することがあります。

図 21

```
apicl# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)

Bonding Mode: fault-tolerance (active-backup)
Primary Slave: None
Currently Active Slave: eth2-1
MII Status: up
MII Polling Interval (ms): 60
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth2-1
MII Status: up
Speed: 10000 Mbps
Duplex: full
Link Failure Count: 1
Permanent HW addr: 6c:ab:05:06:b0:58
Slave queue ID: 0

Slave Interface: eth2-2
MII Status: down
Speed: Unknown
Duplex: Unknown
Link Failure Count: 0
Permanent HW addr: 6c:ab:05:06:b0:59
Slave queue ID: 0
apicl#
```

次のコマンドを実行して、ケーブル接続を検証します。

図 22

```
apicl# acidiag bond0test
Ping test via eth2-1
Setting eth2-2 down
Setting eth2-1 up
Checking whether currently active interface is eth2-1, loopCount: 0
Pinging 10.0.0.30 via bond0/eth2-1
PING 10.0.0.30 (10.0.0.30) 56(84) bytes of data:
64 bytes from 10.0.0.30: icmp_seq=1 ttl=64 time=0.215 ms
64 bytes from 10.0.0.30: icmp_seq=2 ttl=64 time=0.219 ms
64 bytes from 10.0.0.30: icmp_seq=3 ttl=64 time=0.226 ms
64 bytes from 10.0.0.30: icmp_seq=4 ttl=64 time=0.244 ms
64 bytes from 10.0.0.30: icmp_seq=5 ttl=64 time=0.199 ms

--- 10.0.0.30 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4089ms
rtt min/avg/max/mdev = 0.199/0.220/0.244/0.021 ms
Ping via eth2-1 passed
Setting eth2-2 up
Ping test via eth2-2
```

ステップ 2. 新しい APIC からクラスタの状態を確認します。

コンソール、VGA 出力、または SSH で新しい APIC のプロンプトが表示されたら、この APIC のクラスタのビューを調べます。他の APIC サーバーが表示されない場合は、設定パラメータの不一致、ケーブル配線の問題、またはソフトウェア リリースの問題が考えられます。正常な 3 ノード クラスタでは、「acidiag avread」コマンドの出力に、3 台のアクティブなサーバーが表示されます。

図 23

```
apicl# acidiag avread | grep -i active | wc -l
3
apicl# █
```

ステップ 3. データベースの一貫性を確認します。

APIC は、すべての設定およびランタイム データを、シャードと呼ばれるユニットに分割された分散データベースに保存します。シャードは、復元力を確保するためにクラスタ内で 3 重に複製されます。このコマンドを使用すると、データベースが一貫したデータレイヤでクラスタ全体において完全に同期されているかどうかを検査できます。「acidiag rvread」コマンドを使用し、シャードまたはサービス ID マトリックスのどこにもスラッシュまたはバックスラッシュが含まれていないことを確認します。

図 24

