



Subnet Manager の冗長性

この章では、Subnet Manager の冗長性について説明します。内容は次のとおりです。

- [Embedded Subnet Manager \(p.5-3\)](#)
- [High-Performance Subnet Manager \(p.5-4\)](#)
- [マスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ \(p.5-5\)](#)
- [データベースの同期化のセットアップ \(p.5-10\)](#)



(注)

このマニュアルで使用する略語の詳細については、[付録 A 「略語」](#) を参照してください。

Cisco Subnet Manager は、IB 仕様に記載された冗長性をサポートしています。1 台のマスター Subnet Manager と、1 台以上のスタンバイ Subnet Manager で構成されます。マスター Subnet Manager で問題が発生した場合、次のインラインスタンバイ Subnet Manager が IB ファブリックの制御を代行します。

Cisco Subnet Manager には、2 つのタイプがあります。2 つのタイプは、次のとおりです。

- [Embedded Subnet Manager](#)
- [High-Performance Subnet Manager](#)

Embedded Subnet Manager はシャーシ上で動作し、High-Performance Subnet Manager はホスト上で動作します。どちらの Subnet Manager も、相互に IB 標準マスター/スタンバイ フェールオーバーをサポートしています。

Cisco Subnet Manager には、マスター Subnet Manager と 1 台以上のスタンバイ Subnet Manager 間で重要なデータを同期化する、独自のデータベース同期プロトコルが用意されています。これにより、ハイアベイラビリティ冗長が可能になり、IB ファブリックの動作を中断することなく、データベースが同期化されたスタンバイ Subnet Manager がマスターとして制御を代行します。

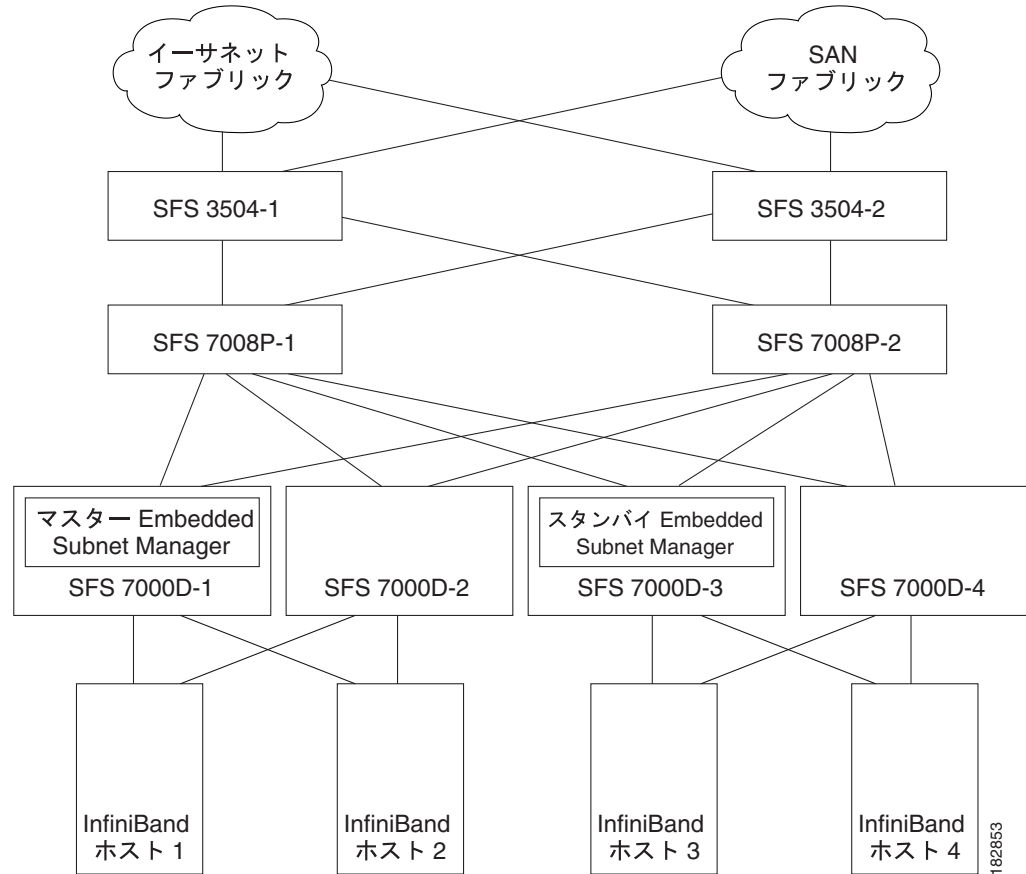


(注)

冗長性のため、2 台の Embedded Subnet Manager を一緒に動作させるか、または 2 台の High-Performance Subnet Manager を一緒に動作させることを推奨します。Embedded Subnet Manager と High-Performance Subnet Manager の間ではデータベースの同期化はサポートされていないため、1 台の Embedded Subnet Manager と 1 台の High-Performance Subnet Manager を設定した場合、障害時にデータトラフィックの中断が発生することがあります。データベースの同期化の詳細については、「[データベースの同期化のセットアップ \(p.5-10\)](#)」を参照してください。

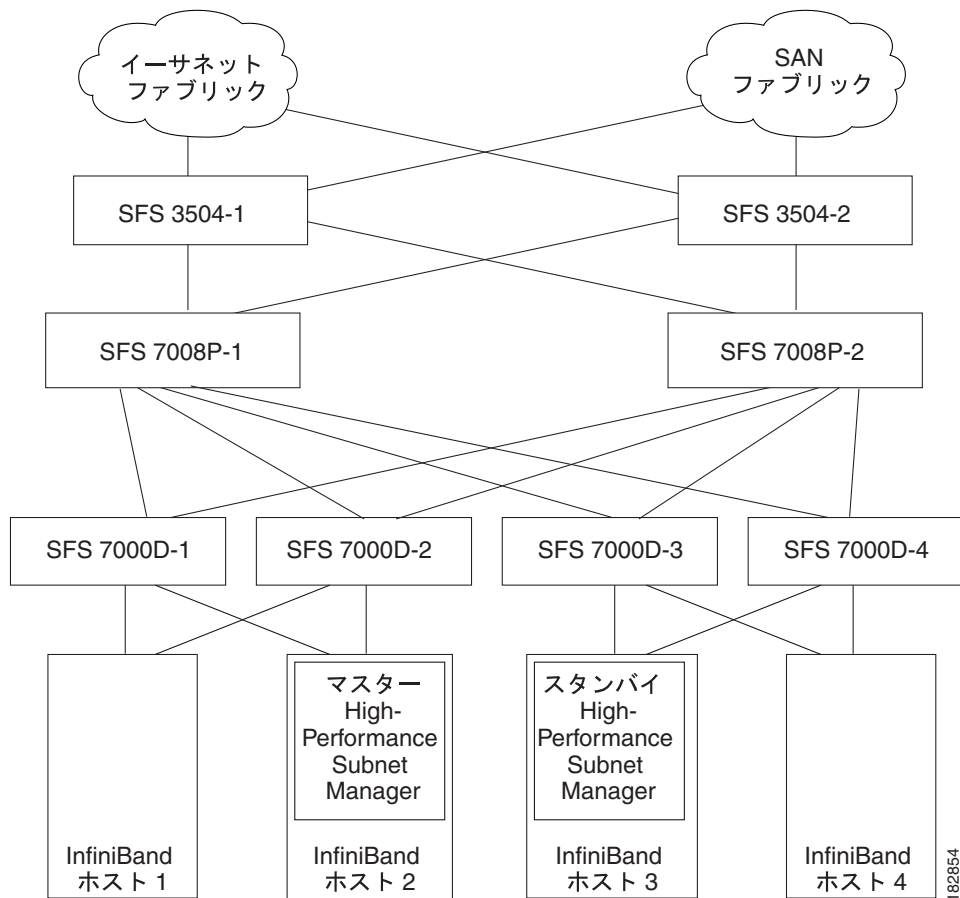
図 5-1 に、Embedded Subnet Manager を使用した IB ファブリックを示します。図 5-2 に、High-Performance Subnet Manager を使用した IB ファブリックを示します。図 5-1 と図 5-2 に示すのは、一般的な Subnet Manager 構成です。用途に応じて、異なる構成が可能です。

図 5-1 InfiniBand ファブリックおよび Embedded Subnet Manager



182853

図 5-2 InfiniBand ファブリックおよび High-Performance Subnet Manager



High-Performance Subnet Manager の詳細については、『Cisco High-Performance Subnet Manager for InfiniBand Server Switches』を参照してください。

Embedded Subnet Manager

ここでは、Embedded Subnet Manager について説明します。

Embedded Subnet Manager は、Cisco SFS 3504、Cisco SFS 3000 シリーズ、Cisco SFS 7000 シリーズ、および Cisco SFS 7008P サーバスイッチ上で動作します。Embedded Subnet Manager をペアで配置すると、システム レベルでシングル ポイント障害を防止できます。Embedded Subnet Manager は、1,000 ノードまでのサブネットでの使用を推奨します。Embedded Subnet Manager を使用すると、大規模サブネット内の変化を、短時間で検出できます。



(注)

Cisco SFS 3012R サーバスイッチおよび Cisco SFS 7008P サーバスイッチは、それぞれのコントローラカード上で Subnet Manager が動作しています。各シャーシには、2 つの Subnet Manager があります。

High-Performance Subnet Manager

ここでは、High-Performance Subnet Manager について説明します。

High-Performance Subnet Manager は、IB サブネットを中央集中的に管理し、制御するスタンドアロンのソフトウェア パッケージです。N+1 コンフィギュレーションとして設定するとハイ アベイラビリティが実現できます。また、大規模クラスタに対する性能とスケーラビリティを向上させることもできます。

ノード数が 1,000 を超えるファブリックでは、High-Performance Subnet Manager の使用を推奨します。Embedded Subnet Manager は Cisco SFS 3504、Cisco SFS 3000 シリーズ、Cisco SFS 7000 シリーズ、および Cisco SFS 7008P サーバ スイッチ上で動作可能ですが、大規模ファブリックでは、メモリ容量が大きく、CPU 性能の高速な High-Performance Subnet Manager がより効率的に機能します。

Cisco High-Performance Subnet Manager は、IB スイッチの組み込みプロセッサから Subnet Manager 機能の負荷を除外することにより、Embedded Subnet Manager を補完します。

High-Performance Subnet Manager は、シャーシ タイプに Embedded Subnet Manager が含まれていない IB ファブリックのネットワーキング構成でも必要です。



(注)

Subnet Manager は、最適なルーティング決定を維持します。スイッチが故障した場合、インバンドトラップ メカニズムによって Subnet Manager への通知が行われます。Subnet Manager は、サブネットのルーティング計算の再実行とルートの再プログラミングにより、リセットを行います。

マスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ

ここでは、マスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ方法について説明します。内容は次のとおりです。

- [Embedded Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ \(p.5-5\)](#)
- [High-Performance Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ \(p.5-7\)](#)



(注)

次のセクションで使用している値は、例として示しています。デフォルト以外の値の使用は推奨しません。

Embedded Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ

ここでは、Embedded Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ方法について説明します。

通常、Embedded Subnet Manager は IB ファブリックの 2 台のシャーシを使用してセットアップし、その他のシャーシではすべてディセーブルにする必要があります。Subnet Manager のプライオリティ番号により、マスターが決定されます。ネットワーク内のすべての Subnet Manager のプライオリティを等しく設定することを推奨します。こうすることにより、ネットワークに新しい Subnet Manager が追加されたときに、ネットワークのマスターシップを引き継がないようにします。

CLI コマンドを使用して Subnet Manager 間のプライオリティの設定および確認を行うには、次の手順を実行します。

- ステップ 1** マスター Subnet Manager およびスタンバイ Subnet Manager を決定するため、Subnet Manager 間のプライオリティを設定します。



(注)

プライオリティの範囲は 0 ~ 15 です。この範囲内で一番大きな値を割り当てられた Subnet Manager が、マスター Subnet Manager となります。デフォルトのプライオリティ番号は 10 です。High-Performance Subnet Manager と Embedded Subnet Manager が同時に同一 IB ファブリック内に存在した場合に、High-Performance Subnet Manager が Embedded Subnet Manager より優先されるように、High-Performance Subnet Manager には Embedded Subnet Manager よりも高いプライオリティが割り当てられます。

次に、2 台の Cisco SFS 3504 サーバスイッチの Embedded Subnet Manager 間にプライオリティを設定する例を示します。

```
SFS-3504# config
SFS-3504(config)# ib sm subnet-prefix fe:80:00:00:00:00:00 priority 12
SFS-3504(config)# exit
```

ステップ 2 master-poll-interval を設定して、マスター Subnet Manager がアクティブかどうかを確認するためのポーリングの間隔を指定します。

次に、master-poll-interval を 5 に設定する例を示します。

```
SFS-3504(config)# ib sm subnet-prefix fe:80:00:00:00:00:00 master-poll-intval 5
```



(注) スタンバイ Subnet Manager による引き継ぎを早くするには master-poll-interval に小さな値を設定し、遅くするには大きな値を設定します。

ステップ 3 master-poll-retries を設定し、マスターへのポーリングの回数を指定します。

次に、master-poll-retries を 0 に設定する例を示します。

```
SFS-3504(config)# ib sm subnet-prefix fe:80:00:00:00:00:00 master-poll-retries 0
```



(注) スタンバイ Subnet Manager による引き継ぎを早くするには master-poll-retries に小さな値を設定し、遅くするには大きな値を設定します。

ステップ 4 設定を確認します。

次に、Subnet Manager の設定、master-poll-interval の値、master-poll-retries の値を確認する例を示します。

```
SFS-3504# show ib sm configuration subnet-prefix all
```

```
=====
Subnet Manager Information
=====
 subnet-prefix : fe:80:00:00:00:00:00
   guid       : 00:05:ad:00:00:01:0c:19
  priority    : 12
   sm-key     : 00:00:00:00:00:00:00:00
 oper-status  : master
   act-count  : 12938
 sweep-interval(sec) : 10
 response-timeout(msec) : 200
 master-poll-intval(sec) : 5
 master-poll-retries : 0
 max-active-sms : 0
 LID-mask-control : 0
 switch-life-time : 18
 switch-hoq-life-time : 18
 host-hoq-life-time : 18
   max-hops    : 64
   mad-retries : 5
 node-timeout(sec) : 10
 wait-report-response : false
 sa-mad-queue-depth : 256
 qos-admin-state : disabled
 max-operational-v1 : auto-link
 min-v1-cap-detected : v10-v17
```

この例の出力では、この Subnet Manager の動作ステータスがマスターで、master-poll-interval が 5 秒で、master-poll-retries が 0 であることが確認できます。

ステップ 5 IB ファブリック内に Subnet Manager があることを確認します。

次に、IB ファブリック内に Subnet Manager が存在することを確認するための `show ib sm sm-info subnet-prefix` コマンドの出力例を示します。

```
SFS-3504# show ib sm sm-info subnet-prefix fe:80:00:00:00:00:00
```

```
=====
                        Discovered Subnet Managers in Fabric
=====
subnet-prefix : fe:80:00:00:00:00:00
port-guid    : 00:05:ad:00:00:01:1d:20
priority     : 0
sm-state     : standby
sm-key       : 00:00:00:00:00:00:00
act-count    : 219
```

この例の `sm-state` は、ファブリック内でこのスタンバイ Subnet Manager が検出されたことを示しています。

High-Performance Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ

ここでは、High-Performance Subnet Manager を使用したマスター Subnet Manager およびスタンバイ Subnet Manager のセットアップ方法について説明します。

通常、High-Performance Subnet Manager は IB ファブリックの 2 台のホストを使用してセットアップし、その他のシャーシではすべてディセーブルにする必要があります。High-Performance Subnet Manager のプライオリティ番号により、マスターが決定されます。ネットワーク内のすべての Subnet Manager のプライオリティを等しく設定することを推奨します。こうすることにより、ネットワークに新しい Subnet Manager が追加されたときに、ネットワークのマスターシップを引き継がないようにします。

CLI コマンドを使用して Subnet Manager 間のプライオリティの設定および確認を行うには、次の手順を実行します。

ステップ 1 マスター Subnet Manager およびスタンバイ Subnet Manager を決定するため、High-Performance Subnet Manager 間のプライオリティを設定します。



(注) プライオリティの範囲は 0 ~ 15 です。この範囲内で一番大きな値を割り当てられた Subnet Manager が、マスター Subnet Manager となります。デフォルトのプライオリティ番号は 11 です。High-Performance Subnet Manager と Embedded Subnet Manager が同時に同一 IB ファブリック内に存在した場合に、High-Performance Subnet Manager が Embedded Subnet Manager より優先されるように、High-Performance Subnet Manager には Embedded Subnet Manager よりも高いプライオリティが割り当てられます。

次に、2 台のホストの High-Performance Subnet Manager 間にプライオリティを設定する例を示します。

```
ib_sm> config priority 12
```

ステップ 2 master-poll-interval を設定して、マスター Subnet Manager がアクティブかどうかを確認するためのポーリングの間隔を指定します。

次に、master-poll-interval を 5 に設定する例を示します。

```
ib_sm> config master-poll-interval 5
```

ステップ 3 master-poll-retries を設定し、マスターへのポーリングの回数を指定します。

次に、master-poll-retries を 0 に設定する例を示します。

```
ib_sm> config master-poll-retries 0
```

ステップ 4 設定を確認します。

次に、Subnet Manager の設定、master-poll-interval の値、master-poll-retries の値を確認するための **show config** コマンドの出力例を示します。

```
ib_sm> show config
```

```
=====
                        Subnet Manager Configuration
=====
      subnet-prefix : fe:80:00:00:00:00:00
          guid      : 00:05:ad:00:00:01:0c:19
      priority      : 12
          sm-key     : 00:00:00:00:00:00:00:00
      oper-status   : master
          act-count  : 2923
      sweep-interval(sec) : 10
      response-timeout(msec) : 200
          mad-retries : 5
          node-timeout : 10
      master-poll-interval(sec) : 5
          master-poll-retries : 0
          max-active-sms : 0
      LID-mask-control : 0
          switch-life-time : 18
          sw-link-hoqlife : 18
          ca-link-hoqlife : 18
          max-hops : 64
      wait-report-response : false
          sa-mad-queue-depth : 256
          local-node-retries : 10
          qos-admin-state : disabled
          max-operational-vl : default
          min-vl-cap-detected : vl0-vl7
ib_sm>
```

この例の出力では、この Subnet Manager の動作ステータスがマスターで、master-poll-interval が 5 秒で、master-poll-retries が 0 であることが確認できます。

ステップ 5 IB ファブリック上の High-Performance Subnet Manager を確認します。

次に、IB ファブリック上の Subnet Manager を確認するための **show other-sm** コマンドの出力例を示します。

```
ib_sm> show other-sm
```

```
=====
                        Subnet Managers in the subnet
=====
 subnet-prefix : fe:80:00:00:00:00:00:00
  port-guid   : 00:05:ad:00:00:01:1d:20
    sm-key    : 00:00:00:00:00:00:00:00
  priority   : 0
  sm-state   : standby
  act-count  : 1133
```

```
ib_sm>
```

この例は、ファブリック内で 1 台のスタンバイ Subnet Manager が検出されたことを示しています。

データベースの同期化のセットアップ

ここでは、データベースの同期化のセットアップ方法について説明します。内容は次のとおりです。

- [Embedded Subnet Manager のデータベースの同期化のセットアップ \(p.5-10\)](#)
- [High-Performance Subnet Manager のデータベースの同期化のセットアップ \(p.5-11\)](#)

Cisco Subnet Manager には、マスター Subnet Manager と 1 台以上のスタンバイ Subnet Manager 間で重要なデータを同期化するための、独自のデータベース同期プロトコルが用意されています。これにより、ハイ アベイラビリティ冗長が可能になり、IB ファブリックの動作を中断することなく、データベースが同期化されたスタンバイ Subnet Manager がマスターとして制御を代行します。これは、企業クラスの IB ファブリックを持ち、MTBF を最小限に抑える必要のある大規模クラスタにおいて重要です。



(注)

ネットワーク内のすべての Subnet Manager のプライオリティを等しく設定することを推奨します。こうすることにより、ネットワークに新しい Subnet Manager が追加された場合、スタンバイとして追加され、マスターに同期化されます。

Embedded Subnet Manager のデータベースの同期化のセットアップ

Embedded Subnet Manager のデータベースの同期化設定をセットアップするには、次の手順を実行します。

ステップ 1 データベースの同期化がイネーブルである (`enable : true`) ことを確認し、現在の設定を表示します。



(注)

デフォルトでは、データベースの同期化機能はイネーブルです。

次に、`show ib db-sync subnet-prefix` コマンドの出力例を示します。

```
SFS-3504# show ib sm db-sync subnet-prefix fe:80:00:00:00:00:00
```

```
=====
Subnet Manager Database Synchronization Information
=====
 subnet-prefix : fe:80:00:00:00:00:00
   enable      : true
 max-dbsync-sms : 1
 session-timeout(sec) : 10
 poll-interval(sec) : 3
 cold-sync-timeout(sec) : 10
   cold-sync-limit : 2
 cold-sync-period(sec) : 900
 new-session-delay(sec) : 120
 resync-interval(sec) : 3600
   state       : in-sync
SFS-3504#
```

ステップ 2 (任意) *max-dbsync-sms* を設定し、マスター Subnet Manager データベースが同期化できるスタンバイ Subnet Manager の最大数を指定します。

次に、*max-dbsync-sms* を 2 に設定する例を示します。

```
SFS-3504(config)# ib sm db-sync subnet-prefix fe:80:00:00:00:00:00 max-dbsync-sms 2
```

ステップ 1 で Subnet Manager Database Synchronization Information の下に表示されるその他の値も、同様にユーザが設定できます。

ステップ 3 Subnet Manager が同期化されていることを確認します。

次に、スタンバイ Subnet Manager を表示するための **show ib sm db-sync subnet-prefix** コマンドの出力例を示します。

```
SFS-3504# show ib sm db-sync subnet-prefix fe:80:00:00:00:00:00 sm-list
```

```
=====
                        DB Synchronizing SMs
=====
      subnet-prefix : fe:80:00:00:00:00:00
      port-guid    : 00:05:ad:00:00:01:1d:20
      entry-state  : active
      session-state : active
      session-timeout-current(sec) : 8
      poll-interval-current(sec)   : 1
      new-session-delay-current(sec) : 120
      resync-interval-current(sec) : 3589
      state        : in-sync
```

```
SFS-3504#
```

この表示では、1 台のスタンバイ Subnet Manager が検出されたことが確認できます。

High-Performance Subnet Manager のデータベースの同期化のセットアップ

High-Performance Subnet Manager のデータベースの同期化設定をセットアップするには、次の手順を実行します。

ステップ 1 データベースの同期化がイネーブルである (*admin-state : enabled*) ことを確認し、現在の設定を表示します。



(注) デフォルトでは、データベースの同期化機能はイネーブルです。

次に、**show db-sync** コマンドの出力例を示します。

```
ib_sm> show db-sync

=====
                        DB Sync Configuration and Status
=====
      protocol-version : 10
        admin-state   : enabled
          max-dbsync-sms : 1
    session-timeout(sec) : 10
      poll-interval(sec) : 3
cold-sync-timeout(sec) : 10
  cold-sync-limit     : 2
cold-sync-period(sec) : 900
new-session-delay(sec) : 120
  resync-interval(sec) : 3600
                        state : in-sync

ib_sm>
```

ステップ 2 (任意) *max-dbsync-sms* を設定し、マスター Subnet Manager データベースが同期化できるスタンバイ Subnet Manager の最大数を指定します。

次に、*max-dbsync-sms* を 2 に設定する例を示します。

```
ib_sm> config db-sync max-dbsync-sms 2
```

ステップ 1 で Database Synchronization Configuration and Status の下に表示されるその他の値も、同様にユーザが設定できます。

ステップ 3 Subnet Manager が同期化されていることを確認します。

次に、スタンバイ Subnet Manager を表示するための **show db-sync sm-list** コマンドの出力例を示します。

```
ib_sm> show db-sync sm-list

=====
                        DB Synchronizing SMs
=====
      port-guid : 00:05:ad:00:00:01:1d:20
        entry-state : active
          session-state : active
    session-timeout-current(sec) : 8
      poll-interval-current(sec) : 1
new-session-delay-current(sec) : 120
  resync-interval-current(sec) : 3373
                        state : in-sync

ib_sm>
```

この表示では、1 台のスタンバイ Subnet Manager が検出されたことが確認できます。