

ユーザー マニュアル- 日本語



PRIMERGY

InfiniBand Host Channel Adapter (100Gb) EDR V1.0

User Manual

(PY-HC331/PY-HC332)

2019年4月版 CA92344-2881-01

■ 商標および著作権について

Linuxは、Linus Torvalds氏の米国およびその他の国における登録商標あるいは商標です。

Red HatおよびRed Hatをベースとしたすべての商標とロゴは、米国およびその他の国におけるRed Hat, Inc.の商標または登録商標です。

その他の各製品名は、各社の商標、または登録商標です。

その他の各製品は、各社の著作物です。

目次

1	はじめに	5
1.1	表記規定	5
1.2	対象読者	6
1.3	ドキュメントの概要.....	6
1.4	概要	7
1.5	製品の搭載条件	8
2	IB HCAカード ハードウェア.....	9
2.1	IB HCAカードについて.....	9
2.2	IB HCAカードのGUID ならびにシリアル番号の識別.....	10
2.3	安全について.....	11
2.4	ENERGY STAR.....	16
2.5	CE 準拠.....	16
3	インターフェース.....	17
3.1	InfiniBand インターフェース	17
3.2	PCI Express インターフェース	18
3.3	LED.....	18
4	IB HCAカードの設定.....	19
4.1	システムの要件	19
4.1.1	ハードウェアの要件.....	19
4.1.2	オペレーションシステム	19
4.1.3	ソフトウェアスタック	19
4.2	取り付け前の注意事項	19

4.3	事前取り付け手順.....	19
4.4	ブラケットの交換手順.....	20
4.5	IB HCAカードの取り付け手順.....	23
4.6	ケーブルの取り付け手順.....	24
5	ドライバのインストール.....	26
5.1	概要.....	26
5.2	ハードウェア/ソフトウェアの要件.....	27
5.3	MLNX_OFEDインストール手順.....	27
5.3.1	RHEL7のMLNX_OFEDインストール.....	28
5.4	MLNX_OFEDアンインストール手順.....	30
5.5	エラッタカーネル適用時のMLNX_OFEDインストール.....	30
5.5.1	MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順.....	30
5.5.2	エラッタカーネルを適用した状態で、MLNX_OFEDをインストールする手順.....	37
5.6	インストール後のファームウェアアップデート.....	38
5.7	Subnet Manager.....	38
5.7.1	OpenSMの機能.....	38
5.7.2	OpenSMの起動確認.....	38
5.7.3	MASTERの確認.....	38
6	トラブルシューティング.....	39
6.1	OpenSM のログファイルを確認する (Linux).....	39
6.2	IB HCA カード.....	41
付録 A	45	
PY-HC331	仕様.....	45
PY-HC332	仕様.....	46

1 はじめに

このマニュアルは PRIMERGY InfiniBand Host Channel Adapter (HCA) の基本的使用方法ならびにハードウェア、ソフトウェア、およびIB HCAカードの搭載とインストールについて説明しています。




製品名称	製品型名
IB HCAカード(100Gbps)	PY-HC331 / PYBHC331
Dual port IB HCAカード(100Gbps)	PY-HC332 / PYBHC332

表 1: 製品名

このマニュアルはさらに、Linux 用のMLNX_OFED (Mellanox Open Fabrics Enterprise Distribution for Linux) のインストール方法について説明しています。

1.1 表記規定

このマニュアルでは以下の表記規定を使用しています。

	注意/警告 この記号は、誤った取扱いをすると、生命が危険にさらされたり、システムが破壊されたり、データが失われる可能性があることを示しています。
	この記号は重要な情報とヒントを示しています。
	この記号は記載されている順序で行う必要がある作業です。
<i>斜体文字</i>	コマンド、メニュー項目、ボタン名、オプション、ファイル名、パス名は本文中では斜体文字で示しています。
<変数>	設定値から変更するべき変数を示しています。
固定フォント	システムの出力は固定フォントで表記しています。
準太字 固定フォント	キーボードから入力されるコマンドは準太字の固定フォントで表記しています。
キー記号	キーはキーボード上の実際の表示に従って表記されます。もし大文字を明確に入力すべき場合は、Shift キーが表記されます。例えば、 SHIFT - A は A を意味します。同時に 2 つのキーを押すべき場合は、2 個のキー記号の間にハイフンを記載しています。

1.2 対象読者

このマニュアルは、InfiniBandネットワークを用いたスイッチ システムプラットフォームのセットアップならびにメンテナンスを担当するユーザー、開発者、システム管理者を対象とします。

このマニュアルは InfiniBand® アーキテクチャ仕様に熟知していることを前提としています。

1.3 ドキュメントの概要



マニュアルの PDF ファイルはインターネットから無償でダウンロードすることもできます。インターネット上のオンラインマニュアルを示す概要ページは、PRIMERGYマニュアルには、「*Industry standard servers*」ナビゲーションオプションを使用してアクセスできます。

Fujitsu Japan から製品を購入した場合は、次の URL を使用してください：

<http://jp.fujitsu.com/platform/server/primergy/manual/>

1.4 概要

この文書は Rack Server に設置される InfiniBand Host Channel Adapter (IB HCA カード) 用のユーザーマニュアルです。このマニュアルに記載されているカードには次の機能があります。

- シングル、またはデュアルポート 100Gbps IB HCA カード
- 交換可能なフルハイト、またはロープロファイル ブラケットを添付したPCIカード
- Mellanox ConnectX-5 を搭載
- PCIe インターフェース
 - PCIe x16 バス インターフェース
 - PCIe 3.0 (8GT/s)
 - オートネゴシエート機能
- IBTA Specification 1.3 準拠
- ハードウェアベース輻輳制御
- QSFP コネクタ28

1.5 製品の搭載条件

システム クラス	サブセグメント	システム	最大搭載 枚数	ブラケット LP/FH
ラック サーバ	エントリー	RX2530 M5	2	LP
		RX2540 M5	2	LP
	ハイエンド	RX4770 M5	4	LP/FH
クラウド サーバ		CX2550 M5	2	LP
		CX2560 M5	2	LP
		CX2570 M5	1	LP

表 2: PRIMERGYの最大搭載数

LP = ロープロファイル ブラケット

FH =フルハイト ブラケット

2 IB HCA カード ハードウェア

2.1 IB HCA カードについて

下表はこのマニュアルで説明するIB HCA カードについて記載しています。

製品名称	PCI Express SERDES Speed	データ伝送 速度/ポート数	ロープロファイル/ フルハイト ブラケット	RoHS
IB HCAカード (100Gbps)	PCIe Gen3 8.0 GT/s	100Gb/s シングルポート	・ロープロファイル ・フルハイト	R-6
Dual port IB HCAカード (100Gbps)	PCIe Gen3 8.0 GT/s	100Gb/s デュアルポート	・ロープロファイル ・フルハイト	R-6

表 3: シングル、デュアルポートカード



図 1: シングルポートカード 部品面



図 2: デュアルポートカード 部品面

2.2 IB HCAカードのGUIDならびにシリアル番号の識別

IB HCAカードには、カードのシリアル番号ならびに InfiniBand プロトコル用のカード GUIDを記載したラベルがIB HCAカードの半田面に貼られています。

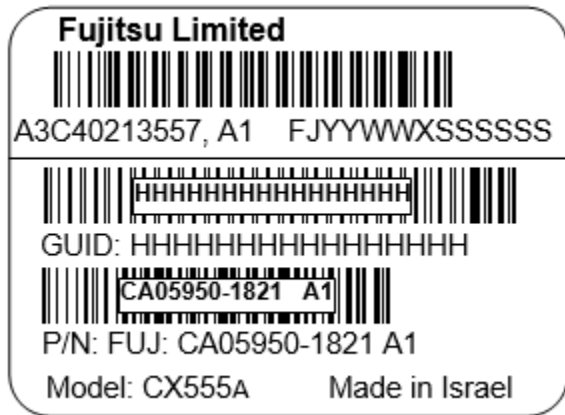


図 3: カード製品ラベル

2.3 安全について



以下の安全上についての注意事項は、『安全上の注意およびその他の重要情報』マニュアルにも記載されています。

このデバイスは IT 機器関連の安全規制に合格しています。目的の環境にサーバを設置できるかどうかについてご質問がある場合は、販売店または弊社カスタマサービス部門にお問い合わせください。



注意!

- このマニュアルに記載されている作業は、技術担当者が行うものとします。技術担当者とは、ハードウェアおよびソフトウェアを含め、サーバを設置するための訓練を受けている要員のことです。
- CSS 障害に関係のないデバイスの修理は、サービス要員が行うものとします。許可されていない作業をシステムに対して行った場合は、保証は無効となり、メーカーの責任は免除されませんので、ご注意ください。
- このマニュアルのガイドラインを遵守しなかったり、不適切な修理を行うと、ユーザーが危険（感電、エネルギーハザード、火災）にさらされたり、装置が破損する可能性があります。
- サーバで内部オプションの取付け、取外しを行う前に、サーバ、すべての周辺装置、および接続されているその他すべてのデバイスの電源を切ってください。また、電源ケーブルをすべてコンセントから抜いてください。ケーブルを抜かなかつた場合、感電の恐れがあります。

作業を始める前に



注意!

- デバイスを設置する際、および操作する前に、お使いのデバイスの環境条件についての指示を守ってください。
- 装置を低温環境から移動した場合は、マシンの内部/外部の両方で結露が発生することがあります。
装置が室温に順応し、完全に乾燥した状態になってから、作業を始めてください。この要件が満たされないと、サーバが破損する場合があります。
- サーバを輸送する際は、必ず元の梱包材に入れるか、あるいは、衝撃からサーバを保護するように梱包してください。

設置と操作



注意!

- このユニットは、必ず周辺温度 35 °C 以下で動作させてください。
- IEC309 コネクタ付き工業用(一般用)電源回路網から電力を供給する設置にこのユニットが組み込まれている場合は、(一般用)電源ユニットのヒューズは、A 型コネクタの非工業用(一般用)電源回路網の要件に準拠している必要があります。
- このユニットは100 V - 240 Vの範囲内で電圧調整されます。主電源がこの範囲内であることを確認してください。
- この装置は正しく接地された電源コンセントまたは、テストされて認可された電源コードでラック内に供給された電源コンセントに接続してください。
- 装置が近くの正しく接地されたコンセントに接続していることを確認してください。



注意!

- 装置の電源ソケットならびに正しく接地された電源コンセントに自由にアクセスできるようにしてください。
- 電源ボタンまたは主電源スイッチ(1つある場合)では、デバイスを主電源電圧から切り離すことはできません。完全に切断するには、すべての電源プラグをコンセントから抜きます。
- サーバと付属の周辺機は常に同じ電源回路に接続してください。さもなければ、仮にサーバがまだ稼働中でも周辺デバイス(例えばメモリ サブシステム)が停電中故障した場合にデータを損失するリスクがあります。
- データケーブルには、干渉を防ぐために適切なシールドを施してください。
- ケーブルは潜在的な危険性が発生せず(躓いたり踏みつけられたりしないこと)、損傷を受けないような方法で敷設してください。サーバを接続する際は、このマニュアルに記載の注意事項を参照してください。
- 荒天時には、データ伝送ケーブルの接続または切断は行わないでください(落雷の危険性

があります)。

- ブレスレットやペーパークリップなどの物や液体がデバイス内部に入る可能性がないことを確認します(感電やショート危険性があります)。
- 緊急時(たとえば、ケーシング、コントロール部分、電源ケーブルの破損や、液体や異物の侵入)には、直ちにデバイスの電源を切り、接地されたコンセントからデバイスを抜き、弊社カスタマサービスセンターに連絡してください。



注意!

- ケーシングが完全に組み立てられ、取り付けベアの背面カバーが所定の位置に取り付けられている(感電、冷却、防火、干渉抑制)場合のみ、(IEC 60950-1 / EN 60950-1 に従って)デバイスの正しい動作が保証されます。
- 安全性、電磁環境適合性、および通信端末機器を規定する要件および規則を満たすシステム拡張機器のみ、取り付けることができます。それ以外の拡張機器を取り付けると、システムが破損したり、安全規定に違反する場合があります。適合するシステム拡張機器についての情報は、弊社カスタマサービスセンターまたは販売店で入手できます。
- 警告ラベル(稲妻マークなど)が付いているコンポーネントを開けたり、取り外したり、交換する作業は、認可された資格を持つ要員以外には行わないでください。例外: CCS コンポーネントの交換は行ってもかまいません。
- システム拡張機器の設置または交換中にデバイスが破損した場合は、保証が無効になります。
- モニタのオペレーティング マニュアルに規定されている解像度とリフレッシュレートのみ設定してください。これを守らなかった場合は、モニタが破損する可能性があります。何かわからないことがございましたら、販売店または弊社カスタマサービスセンターにお問い合わせください。
- 内蔵オプションの取り付け取り外しを行う前に、サーバや、全ての周辺機器、その他接続されている全てのデバイスの電源を切ってください。更にコンセントから電源コードを全て抜いてください。そうしないと感電の危険があります。
- 内部ケーブルまたは装置を損傷または改造しないでください。改造すると装置の故障、火災、または、感電の危険があります。
- サーバ内の装置はシャットダウン後も高温です。シャットダウン後、暫く待ってから内蔵オプションを取り付け、取り外してください。

-
- 回路基板と内蔵オプションの半田付けされた部分は露出していて、露出して、静電気によって損害を受けるおそれがあります。これらを取扱う前に、サーバの金属部分に触って身体から静電気を放電させてください。
 - 基板または、半田付けされた部品の回路に触れないでください。基板の金属部分または端をつかんでください。



注意!

- 内蔵オプションの取り付け、取り外しの時は、元の位置に装置やネジを取り付けてください。異なる種類のネジを使用すると装置が壊れます。
- 本書の設置内容は、オプションにより予告なしに変更することがあります。

バッテリー



注意!

- バッテリーの交換を正しく行わないと、破裂の危険性があります。バッテリーの交換では、まったく同じバッテリーか、またはメーカー推奨バッテリー以外は使用しないでください(システムボードのテクニカルマニュアルを参照)。
- バッテリーはゴミ箱に捨てないでください。バッテリーは特別廃棄物についての自治体の規制に従って、廃棄する必要があります。
- システムボードのテクニカルマニュアルに記載されている指示に従って、システムボードのリチウムバッテリーを交換してください。
- 汚染物質が含まれているバッテリーには、すべてマーク(ゴミ箱の絵に×印)が付いています。また、以下のような汚染物として分類される重金属を示す化学記号が併記されています。

Cd カドミウム

Hg 水銀

Pb 鉛

廃棄処分

この装置の廃棄処分は、全ての国の法律や規制に従う必要があります。

静電気に非常に弱いデバイスが搭載されたモジュール

静電気に非常に弱いデバイスが搭載されたモジュールは、以下のステッカーで識別されます。



図 4: ESD ラベル

ESD ラベル が装着されているコンポーネントを取り扱う際は、必ず以下のポイントを守ってください。

- ESD ラベル が装着されているコンポーネントの取付けや取外しを行う場合は、事前にシステムの電源を切り、コンセントから電源プラグを抜いてください。
- このようなコンポーネントを取り扱う前に、接地された物に触れるなどして静電気の帯電を常に放電してください。
- 使用されるデバイスまたはツールはすべて、非帯電である必要があります。
- 自分とシステムユニットの外部シャーシをつなぐ適切な接地ケーブルを手首に巻いてください。
- ESD ラベル が装着されているコンポーネントを持つ場合は、端または緑色の部分(タッチポイント)を握ってください。
- ESD のコネクタや伝導経路には触れなでください。
- すべてのコンポーネントは、非帯電のパッドの上に置いてください。



ESD コンポーネントの取り扱い方法についての詳細は、欧州規格および国際規格(EN 61340-5-1、ANSI/ESD S20.20)を参照してください。

その他の注意事項:

- このオペレーティング マニュアルとその他のドキュメント(テクニカル マニュアルやドキュメント DVD など)は装置の近くに保管してください。他メーカーに機器を譲渡する場合は、すべてのドキュメントを同梱してください。

2.4 ENERGY STAR



ENERGY STARに準拠し証明された製品は、出荷使用に完全に準拠しています。インストールされるソフトウェアまたは BIOS やエネルギー オプションの変更により、エネルギー消費量は影響されることにご注意ください。そのような場合、ENERGY STAR により保証された特性は保証されません。

「ServerView Operations Manager」ユーザー ガイドには、現在のエネルギー消費量ならびに気温に関するものを含め測定値の読み取り用注意事項が記載されています。Performance Monitor または Task Manager は、CPU 利用レベルを読み取る事に使用できません。



この製品の特定モデルは米国EPAにより認証され、登録されています。これに基づくモデルは、日本市場では販売されていません。

2.5 CE 準拠



システムは、「電磁環境適合性」および2006/95/EC「低電圧指令」についての EC 指令 2004/108/EC の要件に適合しています。このことは、CE マーク(CE = Communauté Européenne)で示されます。

3 インターフェース

IB HCA カードは次のインターフェースが実装されています。

- QSFP28 ポート
- PCI Express x16 レーン
- I/O LEDs

3.1 InfiniBand インターフェース

ConnectX®-5のIB HCA カード はInfiniBand アーキテクチャ仕様 1.3に準拠し、QSFP28 コネクタによりポートにアクセスします。

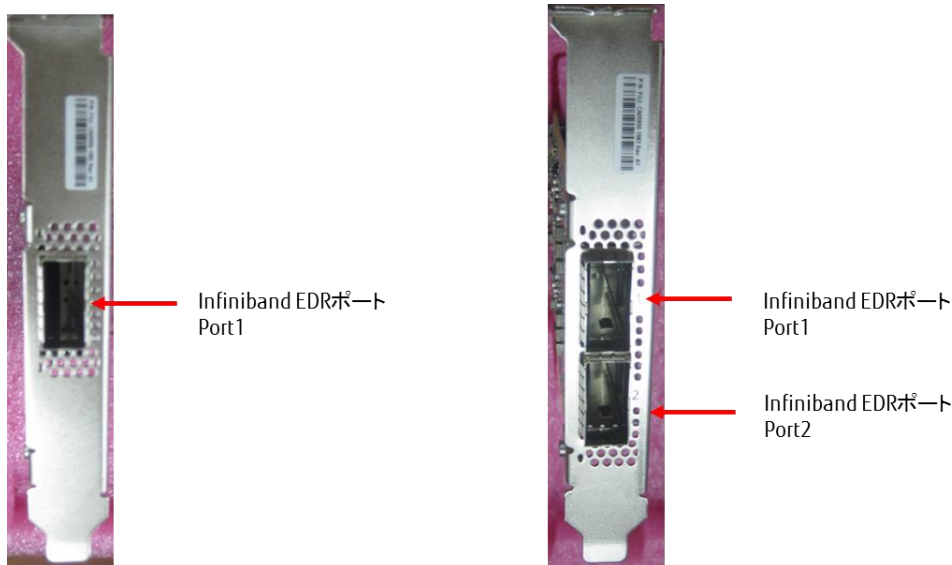


図 5: IB HCAカードのインターフェース(ポート番号)

3.2 PCI Express インターフェース

本製品はPCI Express 3.0 (2.0, 1.1 互換) をサポートし、デバイスは PCI Express バス操作を開始するマスターまたは、PCI バス操作に応答するスレーブのどちらかになります。

3.3 LED

本製品のLEDは、下表の状態を表示します。LEDの位置については「図 6: IB HCAカードのLED」を参照してください。

LED	状態	内容
消灯		物理リンクが確立していない状態
黄色 (イエロー)	点灯	物理リンクがリンクアップしている状態
	点滅	物理リンクがエラーしている状態
緑 (グリーン)	点灯	論理リンクが確立している状態
	点滅	論理リンク上でデータ疎通が行われている状態

表 4: LED表示

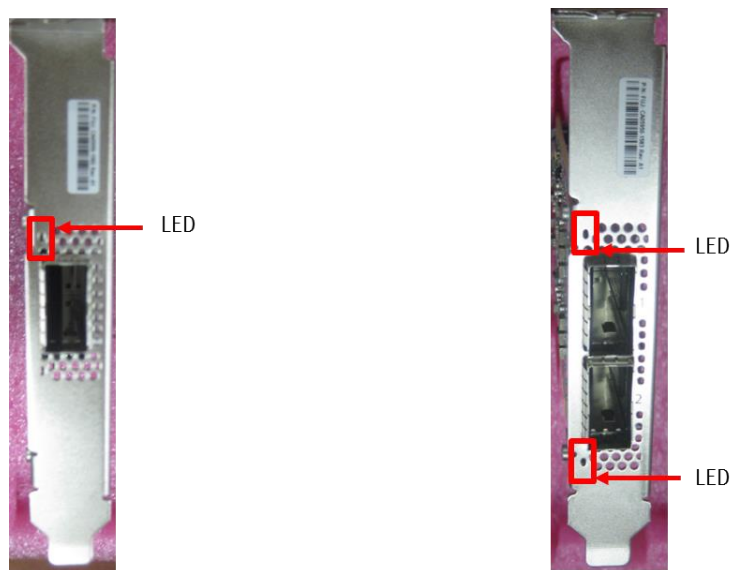


図 6: IB HCAカードのLED

4 IB HCA カードの設定

4.1 システムの要件

4.1.1 ハードウェアの要件

PRIMERGYのPCI Express x16レーンのスロットに本製品を実装してください。

4.1.2 オペレーションシステム

対応するオペレーションシステムは5.2章を参照ください。

4.1.3 ソフトウェアスタック

Mellanox OFED(Linux)のインストール手順については5.3章を参照してください。

4.2 取り付け前の注意事項

1. 身に着けている金属物を外し、本製品を取り扱ってください。
2. MLNX_OFEDを確認してください。
3. PRIMERGYの電源が切れている状態で本製品の取りは外しを行ってください。
4. 作業者は静電ストラップを着用し、作業を行ってください。

4.3 事前取り付け手順

1. インストールするPRIMERGYのPCI Express x16レーンスロット位置とオペレーションシステムを確認してください。
2. PRIMERGYをシャットダウンさせてください。
3. シャットダウン後、PRIMERGYの電源ケーブルを取り外してください。
4. PRIMERGYのPCI Express x16レーンスロットに本製品を取り付けてください。

4.4 ブラケットの交換手順

本製品の出荷時はフルハイトブラケットが取り付けられています。

ロープロファイルブラケットへの交換手順は次の通りです。

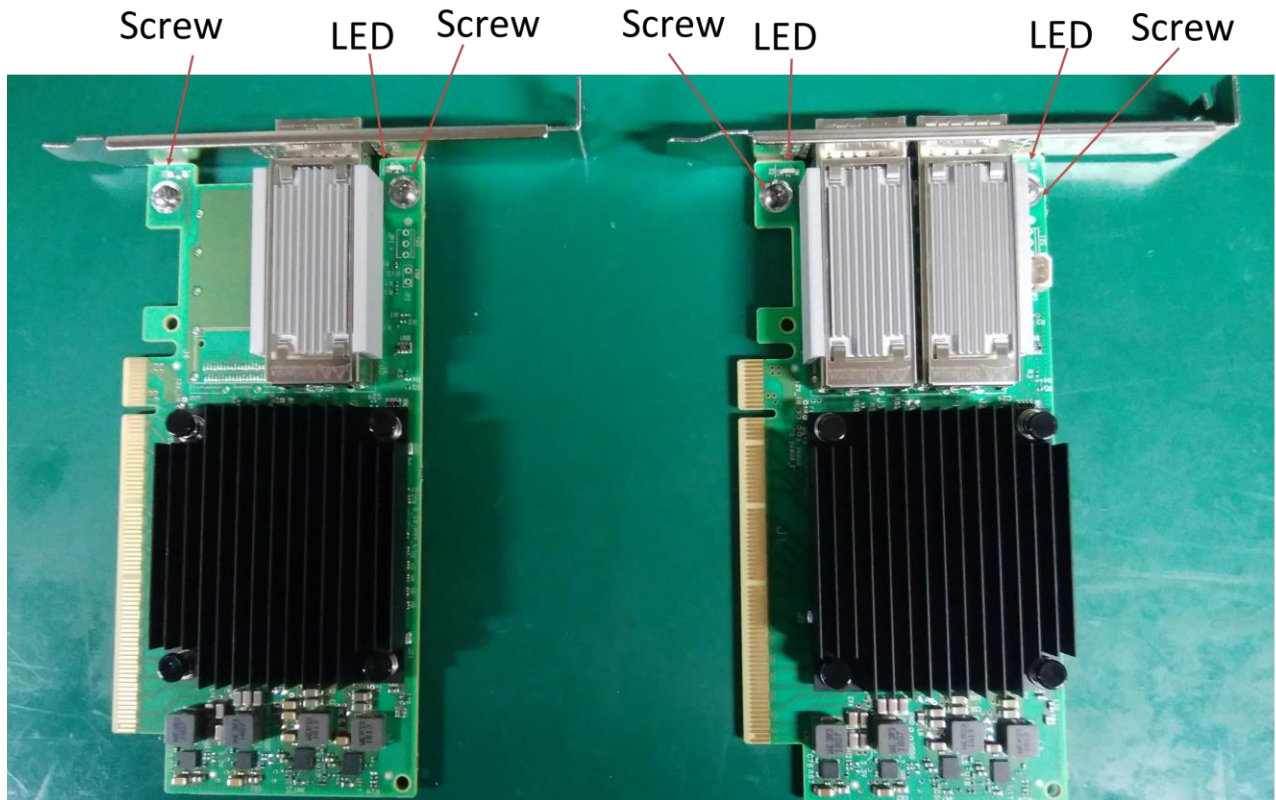


図 7: IB HCAカードの部品名称



図 8: ロープロファイルブラケット

1. ブラケットを固定しているネジ(2本)を取り外します。
2. ブラケットをカードから取り外します。
3. 取り外したネジ(2本)とワッシャー(2個)は紛失しないようにしてください。
4. ネジ穴に合うように、カードの上にブラケットを置いてください。



警告

ブラケットを無理やりカードに押し込まないようにしてください。

5. 取り外していたワッシャーを挟んでネジを締めて、ブラケットをカードに固定してください。
6. ブラケットの LED の穴に、カードの LED の位置が合っているか確認してください

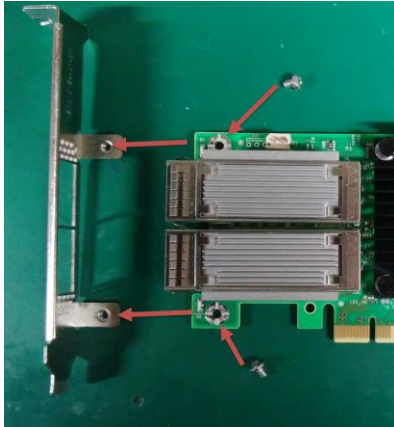


図 9: デュアルポートカードの取付け方法

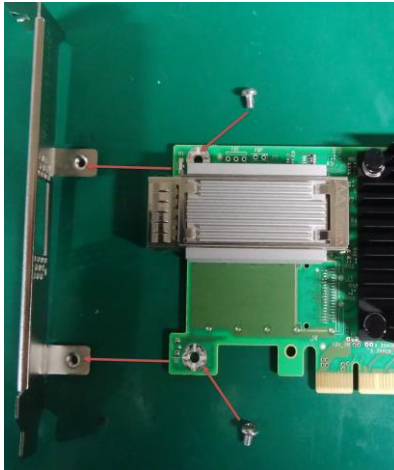


図 10: シングルポートカードの取付け方法

4.5 IB HCAカードの取り付け手順

1. 本製品を取り付ける前に、PRIMERGYがシャットダウンされ、電源ケーブルが接続されていないことを確認してください。作業者は静電気対策し、作業を行ってください。
2. PRIMERGYのカバーを取り外してください。
3. 本製品をPRIMERGYのPCI Express x16レーンスロットに取り付けてください。



警告

本製品はPRIMERGYのPCI-Ex16レーンのスロット以外には取り付けないでください。

4. 本製品の両端をしっかり固定されるまで押し込んでください。



警告

本製品を無理に押し込むと破損する場合がありますので、注意してください。

5. 本製品を、各PRIMERGYの取扱説明書を参照して、取り付けてください。
6. PRIMERGYのカバーを取り付けてください。

4.6 ケーブルの取り付け手順

全てのケーブルは装置の電源が投入された状態で、取付け、取り外しができます。ケーブルを接続する場合、コネクタが確実に接合するまでポートにコネクタを押し込んでください。

LEDは、物理的な接続が完了すると黄色が点灯します(装置の電源が投入されていて、相手側のポートにケーブルが差し込まれている状態で、ケーブルをポートに差し込んだ場合)。ケーブルを接続した後に、ケーブルのロック機構を使ってコネクタをロックしてください。

論理リンクの接続が完了すると、LEDが黄色から緑の点灯に変わります。データが伝送中は、LEDが緑色に点滅します。



警告

ケーブルを取り付ける際は、ラッチのロックを確認してください。



警告

常に、カードと一直線にケーブルとコネクタを押すか、引くことにより、ケーブルを取付け、取外してください。

通気口を通じて排気の流れを妨げない様に注意してください。ケーブルは、ラック内で上下に曲げる前にシャーシの側面を回して、水平に配線できるように長さを考慮してください。

ケーブルを取り外すには、ロックを外して、ポートから丁寧にコネクタを抜き出します。ケーブルが取外されると、LEDが消灯します。



警告

ケーブル(特に長い銅ケーブル)は、相当な重量になるため、IB HCAカード自体にぶらさがらないように、ケーブリングを確認してください。

4.6.1. IB HCAカードへのケーブル取り付け

1. ケーブルをIB HCAカードに取り付ける前に、ケーブルホルダーを使用するか、ケーブルをラックに結んでケーブルの重さを支えてください。
2. コネクタを挿入する前に、カードに対してコネクタの正しい方向を確認してください。間違って上下逆さにコネクタを挿入すると、IB HCAカードを損傷するおそれがあります。
3. コネクタをIB HCAカードに挿入します。コネクタを真直ぐに挿入するように注意してください。IB HCAカードには、上下いずれの向きにも力を加えないでください。
4. コネクタが所定位置でロックされた事を確認してください。

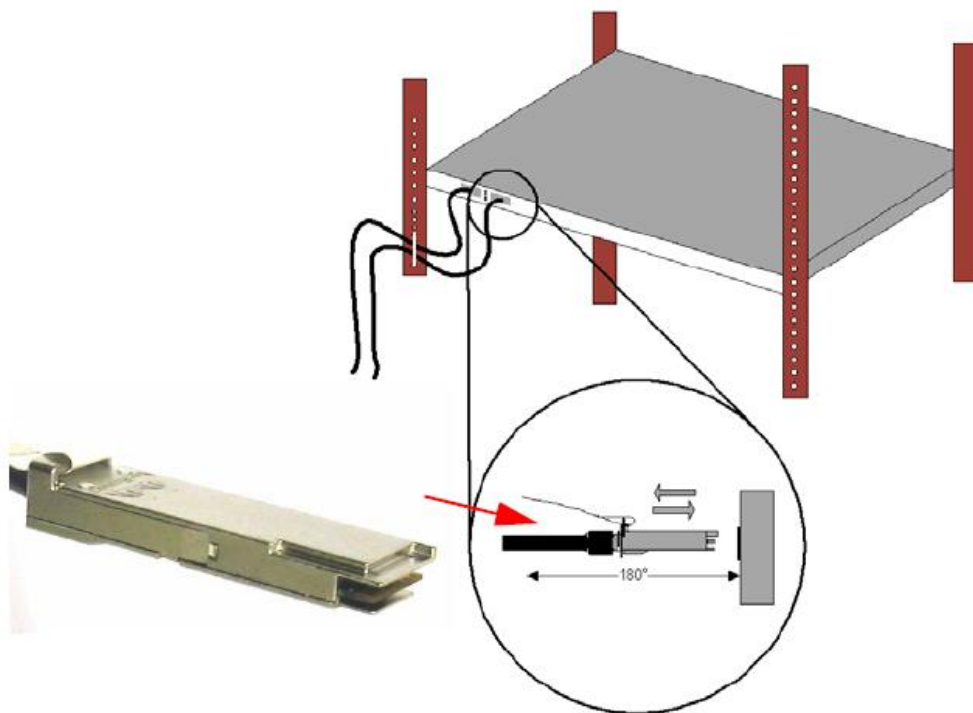


図 11: コネクタの方向

4.6.2. IB HCAカードのケーブル取外し

1. コネクタのロック機構を軽く引っ張ってロックを外し、コネクタをケージから引き出してください。
2. コネクタをIB HCAカードから取り外す時、コネクタをねじらないでください。
3. ケーブルを支えているケーブル支持器具等を取り外してください。

5 ドライバのインストール

5.1 概要

本章では、InfiniBand ハードウェアで構成された単体ホストマシンへのインストール手順について説明します。必要な全てのInfiniBandドライバとソフトウェアは、Red Hat Enterprise Linux (RHEL)のインストール時または、その後にMellanox OpenFabrics Enterprise Distribution for Linux(MLNX_OFED) を使用することでサーバに正しくインストールできます。

最初のインストールには正しいパッケージが選択されている必要があるので、上級者が作業を行ってください。インストールには、MLNX_OFEDドライバパッケージを使用してください。

Fujitsu Japan から製品を購入した場合は、下記よりドライバパッケージがダウンロードできます。

<http://jp.fujitsu.com/platform/server/primergy/downloads/>。



Red Hat Enterprise Linux Version 7.6 以降 をサポートします。

5.2 ハードウェア/ソフトウェアの要件

Linuxへのインストールに必要なディスク容量: 1GB

オペレーティング システム

Linux オペレーティング システム: (x86_64のみ)

インストール権限

インストールには、対象装置の管理者権限が必要です。

5.3 MLNX_OFEDインストール手順

SVIM(Server Installation Manager)でOSをインストールした後、MLNX_OFEDを以下のようにインストールしてください。

Fujitsu Japan から製品を購入した場合は、下記よりドライバパッケージがダウンロードできます。

<http://jp.fujitsu.com/platform/server/primergy/downloads/>



注意

本ソフトウェアはIB HCAカード(100Gps)のドライバパッケージです。

上記からダウンロードしたMLNX_OFEDを使用している場合、OSアップデート時にOS同梱のOFEDを適用しないようにしてください。

5.3.1 RHEL7のMLNX_OFEDインストール

以下はRHEL7.6の例です。

MLNX_OFEDのバージョンにより操作方法が変わる場合がありますので、詳細はMLNX_OFED添付のインストールガイドを参照してください。特に、環境変数(LANG)が日本語ロケールの場合に問題が発生することがあります。

1. ダウンロードしたisoファイルをマウントして、/mntに移動します。
(以下の例では、/mntにマウントしています)

```
[root@localhost /]# mount -ro loop MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-x86_64.iso /mnt/
[root@localhost /]#
[root@localhost /]# cd /mnt/
```

2. 以下のコマンドでインストーラを実行して、*y*を入力します。

```
[root@localhost /]# /mnt/mlnxofedinstall --without-fw-update
Logs dir: /tmp/MLNX_OFED_LINUX-4.5-1.0.1.0.31571.logs
This program will install the MLNX_OFED_LINUX package on your machine.
Note that all other Mellanox, OEM, OFED, or Distribution IB packages will be removed.
Do you want to continue?[y/N]:y

Uninstalling the previous version of MLNX_OFED_LINUX

Starting MLNX_OFED_LINUX-4.5-1.0.1.0 installation ...

Installing mlnx-ofa_kernel RPM
Preparing... #####
mlnx-ofa_kernel #####
Installing kmod-mlnx-ofa_kernel 3.0 RPM
Preparing... #####
kmod-mlnx-ofa_kernel #####
Installing mlnx-ofa_kernel-devel RPM
:
:
:
Installation finished successfully.

Configuring /etc/security/limits.conf.
To load the new driver, run:
/etc/init.d/openibd restart
[root@localhost /]#
```

3. OSを再起動します。

```
[root@localhost /]# reboot
```

4. OS再起動後に、MLNX_OFED版数を確認します。

```
[root@localhost /]# ofed_info | grep -i MLNX_OFED_LINUX  
MLNX_OFED_LINUX-4.5-1.0.1.0 (OFED-4.5-1.0.1.0):
```

5. root アカウントで「`lspci | grep -i Mellanox`」コマンドを実行し、IB HCA カードが認識されていることを確認します。

```
[root@localhost]# lspci | grep -i Mellanox  
02:00.0 Infiniband controller: Mellanox Technologies MT28800 Family  
02:00.1 Infiniband controller: Mellanox Technologies MT28800 Family
```

5.4 MLNX_OFEDアンインストール手順

MLNX_OFEDのアンインストールは以下のコマンドを実行します。

詳細な手順については、「MLNX_OFED添付のインストールガイド」を参照ください。

```
[root@localhost ~]# /usr/sbin/ofed_uninstall.sh
```

その後、OSをリブートします。

```
[root@localhost ~]# reboot
```

MLNX_OFEDをアップグレードする場合は、アンインストールを実行後に新しいOFEDをインストールしてください。

5.5 エラッタカーネル適用時のMLNX_OFEDインストール

エラッタカーネル適用時のMLNX_OFEDのインストール手順について説明します。手順として以下の二種類があります。

- ・MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順
- ・エラッタカーネルを適用した状態で、MLNX_OFEDをインストールする手順

5.5.1 MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順

既にOFEDがインストール済みであり、エラッタカーネル適用のためMLNX_OFEDを再インストールする場合の手順について説明します。インシヤルのMLNX_OFEDインストール手順ではありませんので、ご注意ください。

MLNX_OFED再インストール手順は以下の流れになります。

1. MLNX_OFED のアンインストール
2. エラッタカーネル適用
3. エラッタカーネルに対応した MLNX_OFED の ISO イメージ作成
4. 作成した ISO イメージを使用した MLNX_OFED インストール

1. MLNX_OFED のアンインストール

- 1) MLNX_OFED の ISO イメージを予めダウンロードしておきます。
以下からダウンロード可能です (PRIMERGYダウンロードサイト)

<http://jp.fujitsu.com/platform/server/primergy/downloads/>

- 2) OFED アンインストールの前に `scsi-target-utils` のインストール状況を確認し、インストールされている場合は一旦削除します。`scsi-target-utils` がインストールされていない場合は 3)へ進んで下さい。

```
[root@localhost ~]# rpm -q scsi-target-utils
scsi-target-utils-1.0.24-16.el6.x86_64
```

- 3) ダウンロードしたファイルから ISO イメージを取り出し、以下のように MLNX_OFED のアンインストールを行います。(途中で"y"を入力します)
MLNX_OFEDのアンインストールは、`/usr/sbin/ofed_uninstall.sh`でも出来ます。

MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-x86_64.iso

```
[root@localhost ~]# mount -ro loop MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-x86_64 /mnt
[root@localhost ~]# cd /mnt/
[root@localhost ~]# ./uninstall.sh
```

This program will uninstall all OFED packages on your machine.

```
Do you want to continue?[y/N]:y
Running /usr/sbin/vendor_pre_uninstall.sh
```

Removing OFED Software installations

```
Running /bin/rpm -e mpitests_mvapich_gcc-3.2-923.x86_64
Running /bin/rpm -e mpitests_openmpi_gcc-3.2-923.x86_64
```

```
:
```

```
Stopping IB Subnet Manager...[ OK ]
Running /tmp/6773-ofed_vendor_post_uninstall.sh
[root@localhost ~]#
```

4) OSをリブートします。

```
[root@localhost ~]#reboot
```

2. エラッタカーネル適用

- 1) yum または rpm コマンドでエラッタカーネルの適用を行います。
- 2) OS をリブートします。

3. エラッタカーネルに対応した MLNX_OFED の ISO イメージ作成

- 1) リブート後にカーネル版数を確認します。

```
[root@localhost ~]# uname -r
```

- 2) /mnt 以外のマウント先を作成します。/mnt にマウントすると ISO イメージ作成コマンドが動作しないため、他のマウント先を作成しておきます。以下の例では、/work を作成しています。

```
[root@localhost ~]# mkdir /work
```

- 3) 1. MLNX_OFED のアンインストールで使用した ISO イメージをマウントします。

```
[root@localhost ~]# mount -ro loop MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-x86_64
```

- 4) ISO イメージ作成のコマンドを実行します。(途中で"y"を入力します)
新しいISOイメージは/tmpに作成されます。作成したISOイメージは、/tmpから別のディレクトリへ移動しておくことをお勧めします。

```
[root@localhost ~]# cd /work  
[root@host work]# ./mlnx_add_kernel_support.sh --mlnx_ofed /work/ --make-iso -y  
Note: This program will create MLNX_OFED_LINUX ISO for rhel7.6 under /tmp directory.  
All Mellanox, OEM, OFED, or Distribution IB packages will be removed.  
See log file /tmp/mlnx_ofed_iso.38712.log  
  
Building OFED RPMs. Please wait...  
Running mkisofs...  
Created /tmp/MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-x86_64-ext.iso
```

- 5) /work を umount します。

```
[root@localhost work]# cd ..
```

4. 作成した ISO イメージを使用した MLNX_OFED インストール

1) 作成した ISO イメージを使用して、MLNX_OFED のインストールを行います。(途中で"y"を入力します)

```
[root@localhost ~]# mount -ro loop /tmp/MLNX_OFED_LINUX-4.5-1.0.1.0-rhel7.6-  
x86_64.iso /mnt  
[root@localhost ~]# cd /mnt/  
[root@localhost ~]# ./mlnxofedinstall --without-fw-update  
This program will install the MLNX_OFED_LINUX package on your machine.  
Note that all other Mellanox, OEM, OFED, or Distribution IB packages will be removed.  
  
Do you want to continue?[y/N]:y  
  
Starting MLNX_OFED_LINUX-4.5-1.0.1.0 installation ...  
  
Installing kernel-ib RPM  
準備中... #####  
kernel-ib #####  
:  
:  
Installation finished successfully.  
  
Please reboot your system for the changes to take effect.
```

2) /mnt を umount します。

```
[root@localhost ~]# umount /mnt
```

- 3) 1. MLNX_OFED のアンインストールで `scsi-target-utils` を削除した場合は、再度インストールを行います。なお、インストールするパッケージは Red Hat Enterprise Linux のインストールディスクに収録されています。以下の例では、前記インストールディスクを使用しています。

対象のrpmパッケージ

`scsi-target-utils-1.0.14-4.el6.x86_64.rpm`

```
[root@localhost ~]# mount /dev/cdrom /mnt
mount: ブロックデバイス /dev/sr0 は書き込み禁止です、読み専用でマウントします
[root@localhost ~]# cd /mnt/Packages/
[root@localhost Packages]# rpm -ivh scsi-target-utils-1.0.24-16.el6.x86_64.rpm
警告: scsi-target-utils-1.0.24-16.el6.x86_64.rpm: ヘッダ V3 RSA/SHA256 Signature, key ID
fd431d51: NOKEY
準備中...      ##### [100%]
 1:scsi-target-utils  ##### [100%]
[root@localhost ~]# rpm -q scsi-target-utils
scsi-target-utils-1.0.24-16.el6.x86_64
```

- 4) `/mnt` を `umount` し、インストールディスクを取り出します。

```
[root@localhost Packages]# cd ..
```

- 5) OS をリブートします。

```
[root@localhost ~]# reboot
```

- 6) OS 起動後に `limits.conf` ファイルに以下の 2 行が含まれていることを確認します。

```
[root@localhost ~]# tail -2 /etc/security/limits.conf
* soft memlock unlimited
* hard memlock unlimited
```

7) IBドライバがロードされていることを確認します。

```
[root@localhost ~]# /etc/init.d/openibd status
```

```
HCA driver loaded
```

```
Configured IPoIB devices:
```

```
ib0 ib1
```

```
Currently active IPoIB devices:
```

```
The following OFED modules are loaded:
```

```
rdma_ucm  
rdma_cm  
ib_addr  
ib_ipoib  
mlx4_core  
mlx4_ib  
mlx4_en  
ib_mthca  
ib_uverbs  
ib_umad  
ib_ucm  
ib_sa  
ib_cm  
ib_mad  
ib_core  
iw_nes
```

5.5.2 エラッタカーネルを適用した状態で、MLNX_OFEDをインストールする手順

OSインストールからエラッタカーネルを適用し、MLNX_OFEDをインストールする場合の手順について説明します。イニシャルのMLNX_OFEDインストール手順ではありませんので、ご注意ください。

MLNX_OFEDインストール手順は以下の流れになります。

1. OSのインストール
2. エラッタカーネル適用
3. エラッタカーネルに対応したOFEDのISOイメージ作成
4. 作成したISOイメージを使用したOFEDインストール

1. OSのインストール

OSをインストールします。

2. エラッタカーネル適用

「5.5.1 MLNX_MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順」の「2. エラッタカーネル適用」の手順を実行します。

3. エラッタカーネルに対応したMLNX_OFEDのISOイメージ作成

「5.5.1 MLNX_MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順」の「3. エラッタカーネルに対応したMLNX_OFEDのISOイメージ作成」の手順を実行します。

4. 作成したISOイメージを使用したMLNX_OFEDインストール

「5.5.1 MLNX_MLNX_OFEDインストール状態で、エラッタカーネルを適用する手順」の「4. 作成したISOイメージを使用したMLNX_OFEDインストール」の手順を実行します。

パッケージの削除のメッセージが表示された場合、オプション(--nodeps)を追加し、表示されているパッケージからscsi-target-utilsを除外して、パッケージをアンインストールします。その後、再度インストールを実行します。MLNX_OFEDのバージョンによっては、自動でパッケージがアンインストールされます。

5.6 インストール後のファームウェアアップデート

新しいファームウェアを書き込む場合は、Fujitsu Technology Solutions Webサイトからダウンロードできます (<http://support.ts.fujitsu.com>) > *Drivers & Downloads*)。



Fujitsu Japan から製品を購入した場合は、下記よりダウンロードしてください。

<http://jp.fujitsu.com/platform/server/primergy/downloads/>

5.7 Subnet Manager

少なくとも1つのSubnet Manager (SM)がサブネットに存在します。各SMは、IB HCAカードまたはIBスイッチのポートに常駐します。サブネット上に複数のSMが存在する場合、そのうちの1つがMASTERとなり、その他のSMはSTANDBYとなります。MASTER SMはサブネットの初期化及び設定をする役割を持っています。

OpenSM はInfiniBand準拠のSMです。OpenSMの利用法を以下に説明します。

5.7.1 OpenSMの機能

Open SM は RHEL に含まれている Subnet Manager です。この章では、MASTER SM となるサーバの設定について説明します。以下を参照して MASTER となるサーバを設定してください。OpenSM の他の設定については、`opensmd` のコマンドヘルプを参照してください。

基本的には、最小の GUID を持つサーバがサブネットの MASTER SM となります。MASTER SM を選択する場合、0 から 15 までの優先順位を決めることができます。15 が最高、0 が最低の優先順位になります。

RHEL7 の例: `# /etc/init.d/opensmd start -p 15`

5.7.2 OpenSMの起動確認

次のコマンドを入力して OpenSM が起動していることを確認してください。

```
# /etc/init.d/opensmd status
```

このコマンドの実行結果、"running"が表示されれば、OpenSMは正しく起動しています。

コマンドの実行結果、"stopped"が表示された場合は、次のコマンドを入力して OpenSM を起動させてください。

```
# /etc/init.d/opensmd start
```

5.7.3 MASTERの確認

サーバで `ibstat` コマンドを入力すると、"Base LID"ならびに"SM LID"が表示されます。

上記のLIDと同じLIDである IB HCA カードならびに InfiniBand スイッチが MASTER です。

6 トラブルシューティング

この章では、本製品が正常に動作しない場合について説明しています。

対処後もトラブルが解消しない場合は、リペアセンターに連絡してください

Fujitsu Japan から製品を購入した場合は、お客様担当の弊社システムエンジニアに問合わせください。

6.1 OpenSM のログファイルを確認する (Linux)

MASTER として起動しているサブネットマネージャ (OpenSM) のサーバのログファイル (*/var/log/opensm.log*) を確認します。

チェックリスト: ログファイル (*/var/log/opensm.log*) に、次のエラーメッセージがあるか確認します。

- SM port is down
- Removed port with GUID:GUID number PRIx64 LID [LID number, LID number] of node:
- Reporting Generic Notice type:1 num:128 (Link state change) from LID:LID number GUID:GUID number

SM port is down

サブネットマネージャの MASTERポートの接続が切断されました。

対処方法: 次の交換が必要です。リペアセンターまでご連絡ください。

1. IB HCA カード

MASTERのサブネットマネージャが動作している IB HCA カードを交換してください。

2. システムサーバ

上記の IB HCA カードを交換してもメッセージが表示される場合は、IB HCA カードを搭載しているサーバを交換してください。

3. InfiniBand スイッチ

上記のサーバを交換してもメッセージが表示される場合は、InfiniBand スイッチを交換してください。

Removed port with GUID:GUID number PRlx64 LID [LID number, LID number] of node:

IB HCA カード、InfiniBand スイッチ、InfiniBand ケーブルのいずれかの接続が切断されました。

対処方法：次のユニットのいずれかの交換が必要です。リペア センターに連絡してください。

1. IB HCA カード

MASTERのサブネットマネージャが動作している IB HCAカードを交換してください。エラー メッセージの GUID(またはLID)情報と ibnetdiscover 出力結果から、該当するIB HCA カードを特定します。

エラーメッセージの例：

```
Jul 30 23:34:18 600287 [4CD0D940] 0x02 ->
__osm_drop_mgr_remove_port: Removed port with
GUID:0x0002c90300041306 PRlx64 LID range [2, 2] of node:localhost HCA-2
この場合は、GUID が 0x0002c90300041306 で、LID が 2 のIB HCA カードが該当します。
```

2. InfiniBand スイッチ

該当する InfiniBand スイッチを交換してください。エラーメッセージの GUID(またはLID)情報と ibnetdiscover 出力情報から、該当する InfiniBand スイッチを特定します。

エラーメッセージの例：

```
Jul 30 23:34:18 600424 [4CD0D940] 0x02 ->
__osm_drop_mgr_remove_port: Removed port with
GUID:0x0002c9020040c790 PRlx64 LID range [21, 21] of node:(none) SW-1
この場合は、GUID が 0x0002c9020040c790 で、ID が 21 のInfiniBand スイッチが該当します。
```

3. IB ケーブル

上記の InfiniBand スイッチを交換してもメッセージが表示される場合は、InfiniBand スイッチに接続されているすべての IB ケーブルを交換してください。

6.2 IB HCA カード

サーバで下記の項目を確認して、エラーが発生している箇所や発生時間を確認します。

項	確認アイテム
1	IB HCA カードの物理リンク(イエロー)、データ転送(グリーン)LED
2	IB HCA カードドライバのメッセージ

表 5: IB HCAカード確認アイテム

IB HCA カードの物理リンク(イエロー)、データ転送(グリーン)LED

次の確認内容によりエラーの発生している箇所を確認できます。

項	確認内容	判断内容	対処方法
1	サーバは起動していますか？	はい	2 へ
		いいえ	サーバを起動してください。
2	IB HCAカードは正しく搭載されていますか？	はい	3 へ
		いいえ	IB HCA カードを正しく搭載してください。
3	IBケーブルは接続されていますか？	はい	4 へ
		いいえ	IB ケーブルを IB HCA カードに接続してください。
4	接続されたIB スイッチは正常に動作していますか？	はい	1. IB HCAカードを交換してください。 2. IB HCAカードを交換しても物理リンクLED が消灯している場合は、サーバ本体を交換してください。 3. 更にサーバ交換後も物理リンクLEDが消灯している場合は、IB ケーブルを交換してください。 4. 上記対処後もLEDが点灯しない場合はご連絡ください: technical.support@ts.fujitsu.com Fujitsu Japan から製品をご購入の場合、お客様担当の弊社システムエンジニアに問い合わせください。
		いいえ	IB スイッチの交換が必要です。

表 6: 物理リンクLED (グリーン) が消灯または点滅

項	確認内容	判断内容	対処方法
1	データ転送LED (イエロー) が消灯していますか？	はい	OpenSMのMASTERサーバでOpenSMが動作していることを確認してください。
		いいえ	上記の確認項目でエラーの発生場所を確認してください。(表 5 物理リンクLED が消灯または点滅)

表 7: データ転送LED (イエロー)が消灯

IB HCAカードドライバ メッセージ (Linux)

`/var/log/messages` のログファイルを参照し、InfiniBand カードドライバのエラーメッセージを確認してください。ログファイルの参照は、テキストエディタ等を使用してください。

項	メッセージ内容	対処方法
1	<code>dev_err(&dev->pdev->dev, "%s</code>	IB HCA カードの交換が必要です。 下記に連絡してください。 technical.support@ts.fujitsu.com Fujitsu Japan から製品をご購入の場合、 お客様担当の弊社システムエンジニアに 問い合わせください。
2	<code>dev_err(&dev->pdev->dev, "handling bad device here\n");</code>	
3	<code>dev_err(&dev->pdev->dev, "device's health compromised - reached miss count\n");</code>	
4	<code>dev_err(&dev->pdev->dev, "NIC was disabled\n");</code>	
5	<code>dev_err(&pdev->dev, "Failed initializing command interface, aborting\n");</code>	
6	<code>dev_err(&dev->pdev->dev, "Firmware over %d MS in initializing state, aborting\n",</code>	

表 8: IB HCA カードドライバのエラーメッセージ

項	確認内容	判定内容	対処方法
1	IB HCA カードは正しく実装されていますか？	はい	IB HCA カードの交換が必要です。 IB HCA カードを交換してもメッセージが表示される場合はサーバ本体の交換が必要です。
		いいえ	IB HCA カードを正しく実装し直してください。

表 9: IB HCAカード対処方法

IB HCA カード交換後の確認作業 (Linux)

1. サーバ本体の電源を投入します。
2. Linux を起動して root アカウントで `/sbin/lspci` コマンドを実行し、IB HCA カードが認識されていることを確認します。

実行結果:

```
02:00.0 Infiniband controller: Mellanox Technologies MT28800 Family
```

```
02:00.1 Infiniband controller: Mellanox Technologies MT28800 Family
```

付録 A

PY-HC331 仕様

本体		電源と環境	
サイズ:	2.71in x 5.6in (68.90mm x 142.25mm)	電圧:	12V
風量:	400LFM @3.5W QSFP cable 350LFM @1.5 W QSFP cable	電力(typ):	パッシブ ケーブル 14.2W
QSFP28 100Gb/s コネクタ:	InfiniBand (銅と光) ポートごとの最大電力は 3.5W。	最大電力:	パッシブ ケーブル 17.0 W
		温度:	0°C から 55°C

表 10: シングルポートカードの仕様 1

プロトコル サポート		規格	
InfiniBand:	IBTA v1.3(plan) Auto-Negotiation (100Gb/s, 25Gb/s per lane) または (56Gb/s, 14Gb/s per lane) または (40Gb/s, 10Gb/s per lane) または (20Gb/s, 5Gb/s per lane) または (10Gb/s, 2.5Gb/s per lane)	安全性:	CB/CTUVus/CE
QoS:	各ポートに 8 本の InfiniBand 仮想レーン	RoHS:	RoHS-R6
RDMA サポート:	有り。すべてのポート	EMC	CE/FCC/VCCI/ICES/RCM
Data Rate:	最大 100 Gb/s EDR – InfiniBand		
PCI Express:	SERDES @ 8.0GT/s, 16 レーン		

表 11: シングルポートカードの仕様 2

PY-HC332 仕様

本体		電源と環境	
サイズ:	2.71 in. x 5.6 in. (68.90mm x 142.25 mm)	電圧:	12V, 3.3V
風量:	650LFM @3.5W QSFP cable 350LFM @1.5 W QSFP cable	電力(typ):	パッシブ ケーブル 19.3W
QSFP28 100Gb/s コネクタ:	InfiniBand (銅と光) ポートごとの最大電力は 3.5W。	最大電力:	パッシブ ケーブル 23.4 W
		温度:	0°C から 55°C

表 12: デュアルポートカードの仕様 1

プロトコル サポート		規格	
InfiniBand:	IBTA v1.3(plan), Auto-Negotiation (100Gb/s, 25Gb/s per lane) または (56Gb/s, 14Gb/s per lane) または (40Gb/s, 10Gb/s per lane) または (20Gb/s, 5Gb/s per lane) または (10Gb/s, 2.5Gb/s per lane)	安全性:	CB/cTUVus/CE
QoS:	各ポートに 8 本の InfiniBand 仮想レーン	RoHS:	RoHS-R6
RDMA サポート:	有り。すべてのポート	EMC	CE/FCC/VCCI/ICES/RCM
Data Rate:	最大 100 Gb/s EDR- InfiniBand		
PCI Express:	SERDES @ 8.0GT/s, 16 レーン		

表 13: デュアルポートカードの仕様 2

InfiniBand Host Channel Adapter (100Gb) EDR V1.0
User Manual
(PY-HC331/PY-HC332)

発行日 2019年4月
発行責任 富士通株式会社
CA92344-2881-01

- 本書の内容は、改善のため事前連絡なしに変更することがあります。
- 本書に記載されたデータの使用に起因する、第三者の特許権およびその他の権利の侵害については、当社はその責を負いません。
- 無断転載を禁じます。