

ホワイトペーパー

BladeSymphony Virtage SR-IOV のご紹介

2014 年 7 月発行

株式会社 日立製作所

目次

1	はじめに.....	3
1.1	登録商標・商標.....	3
1.2	注意事項.....	3
1.3	免償事項.....	3
2	SR-IOV の機能概要.....	4
3	性能検証事例.....	5
3.1	測定環境.....	5
3.2	測定結果.....	6
3.3	各方式による共有 NIC 性能比較(ポートあたり).....	6
4	ユースケース(BS2000 での大規模データベースシステム).....	7
5	使用上の注意事項.....	7
	付録 1. ソフトウェア方式による NIC 共有.....	8

1. 本書の内容は一部または全部を無断転載することは禁止されています。
2. 本書の内容に関しては将来予告なしに変更することがあります。
3. 日立製作所の許可なく複製・改変などを行うことはできません。
4. 日立製作所が製品やサービスについて行う保証は、製品添付の保証文章に記載した内容のみに限定され、本書の
どの箇所であっても何ら新規の保証を行うものではありません。
5. 運用した結果の影響については、責任を負いかねますのでご了承ください。
6. 本書に技術的あるいは編集上の誤りや欠陥があったとしても、日立製作所は一切の責任を負わないものとします。

1 はじめに

本書は仮想化環境において利用する SR-IOV (Single Root I/O Virtualization) の機能概要と Virtage 環境下での構成例や性能検証結果について記載しており、SR-IOV の活用シーンについて理解を深めていただくことを目的としています。

1.1 登録商標・商標

- Intel, Xeon はアメリカ合衆国および/またはその他の国における Intel Corporation の商標です。
- 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 は米国およびその他の国における Red Hat, Inc の登録商標あるいは商標です。
- Emulex は、米国 Emulex Corporation の登録商標です。
- その他記載の会社名、製品名はそれぞれの会社の登録商標あるいは商標です。

1.2 注意事項

性能については、ベンチマーク値を使った机上計算のみでなく、実際の運用環境と同等構成(ハード・ソフト)において、要件を満たしていることを確認してください。ベンチマーク値の単純比較のみならず、本番相当環境(実機)を利用した事前評価等を強く推奨します。システム性能は、テスト方法やハードウェアに搭載するプロセッサの種類・動作周波数、メインメモリ・キャッシュメモリ容量、ディスク数・ディスク構成、動作ソフトウェア構成等の条件により大きく左右されます。サイジングなどにおいてはベンチマーク値を目安としてご利用いただけますが、机上計算のみでは思わぬ性能不足等に陥る可能性があります。

1.3 免責事項

本サイトに掲載している性能評価指標またはその他の性能に関連する情報の使用によって、いかなる損害が生じた場合も、情報提供者である日立製作所は責任を負いません。

2 SR-IOV の機能概要

SR-IOV とは PCI デバイス側で仮想化をサポートする機能であり、デバイス共有処理をハードウェア側で実装します。SR-IOV では、1 つの物理 NIC デバイス(PF: Physical Function)において複数の仮想的な NIC デバイス(VF: Virtual Function)を生成し、ハードウェア内に実装されている NIC 内ネットワークスイッチに PF および VF を接続することで NIC デバイスの共有を実現します。

これにより、ソフトウェアによる NIC 共有に比べ、PF と同等の最大帯域を確保できます。また帯域は VF 毎に制限を掛けることができ(100Mbps 単位)、性能確保が必要なポートに対し、リソースを確実に割り当てることが可能です。さらにホスト CPU のデバイス共有処理に関わる負荷をゼロにできるため、広帯域通信が可能となります。ネットワークの高スループットが必要な基幹業務システム等に適しています。以下に Virtage における SR-IOV を使用した NIC 共有の概略図を示します。

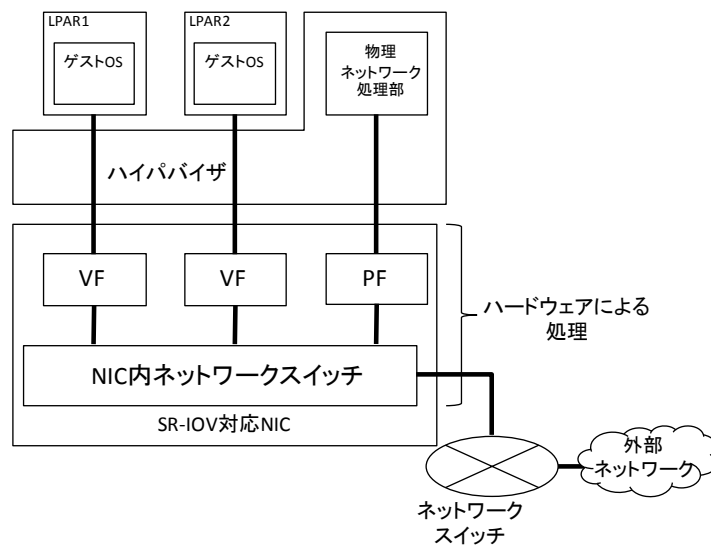


図 2-1 SR-IOV による NIC 共有の概略図

3 性能検証事例

SR-IOV による NIC 共有のスループット性能(Mbps: 単位時間あたりのデータ転送量)を測定した結果を示します。5LPAR 構成と 15LPAR 構成のシステム装置に対し、LPAR 毎に 2+2(待機系)の VF を設定しました。

3.1 測定環境

表 3-1. 測定環境

サーバブレード	BS2000 高性能ブレード A2 モデル	
EFI	07-61	
Virtage	79-70	
OS	Red Hat Enterprise Linux 6.4(x64)	
LPAR	5LPAR	15LPAR
現用 VF/LPAR	2	
現用 VF 合計	10	30
PF	2+2(待機)	
CPU	Intel (R) Xeon (R) CPU E7-8870	
メモリ容量	32GB	
内蔵スイッチ	10Gbps DCB スイッチ	
NIC	管理 LAN	オンボード
	業務 LAN	Emulex 10Gb 2port LAN 拡張カード ×2 枚
	FW	4.6.348.0
	Driver	4.6.352.0

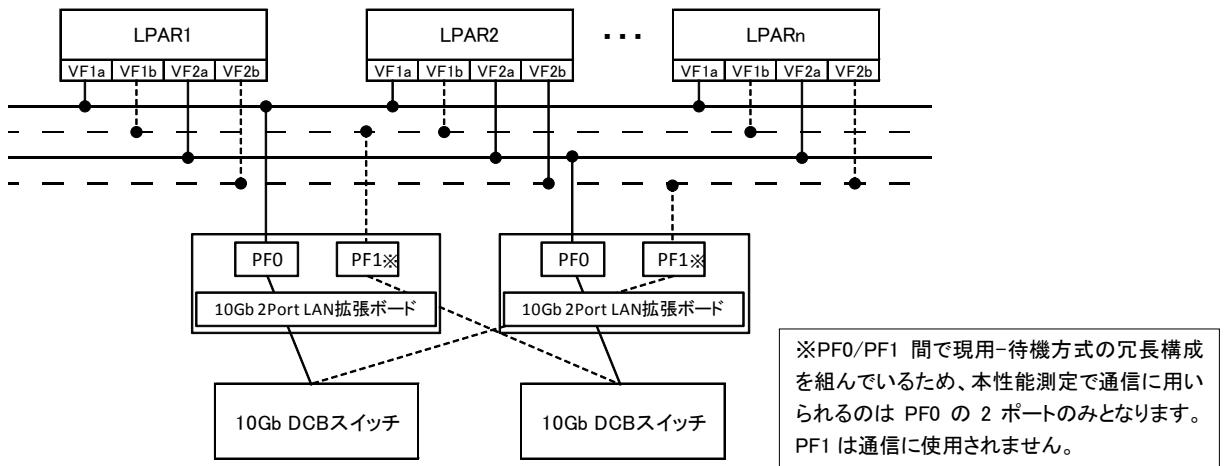


図 3-1 NIC 共有の概略図(SR-IOV)

3.2 測定結果

SR-IOVによるNIC共有では5LPAR、15LPAR共に物理1ポートあたり約8Gbpsの通信性能となりました。また各LPAR毎に大きな偏りなく、物理リソースを均等に分割しております。さらに最大スループット時のホストCPU負荷の大幅な低減を得ることができました。

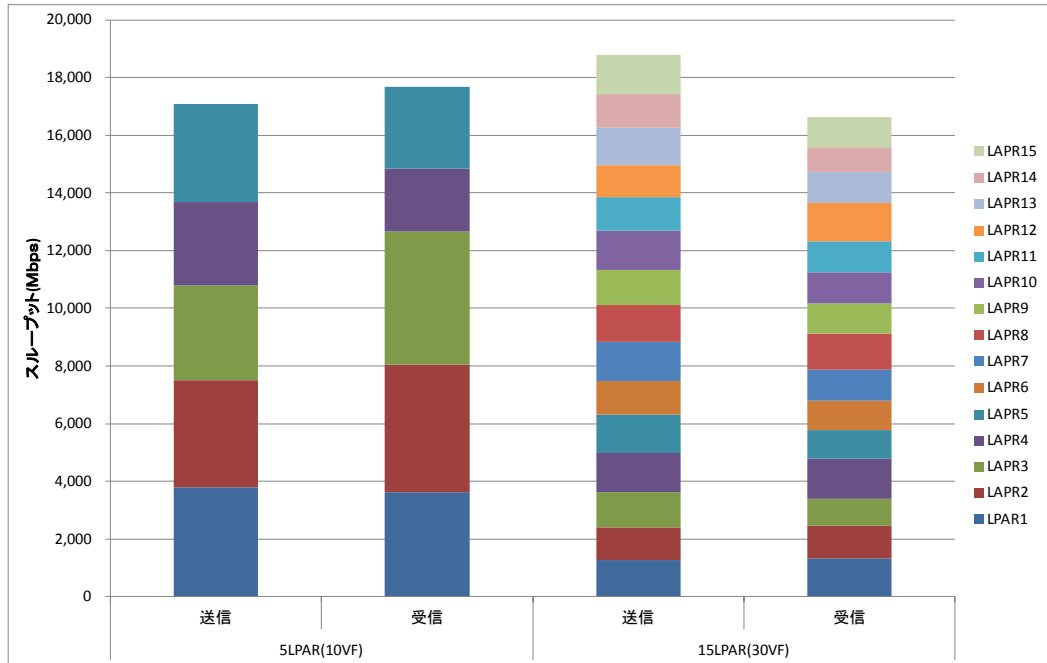


図 3-2 SR-IOV による NIC 共有時の合計スループット*

*グラフ中の一つの LPAR のスループットは現用系の 2 つの物理ポートの合計値を示します。

3.3 各方式による共有 NIC 性能比較(ポートあたり)

ソフトウェアによるNIC共有はソフトウェアによる実装であるため、1Gbpsの通信を行った場合にホストCPUへの負荷が大きく、10GbpsのNICデバイスを使用しても、2個のCPUを常に使用して約2.2Gbpsの性能となります。VirtageのSR-IOVによるNIC共有はハードウェア側でNIC共有処理を実施するため、ホストCPUへの負荷は削減され、物理NICデバイスと同等の約8Gbps(10GbpsのNICカード使用時)までの性能が期待でき、広帯域通信が可能となります。

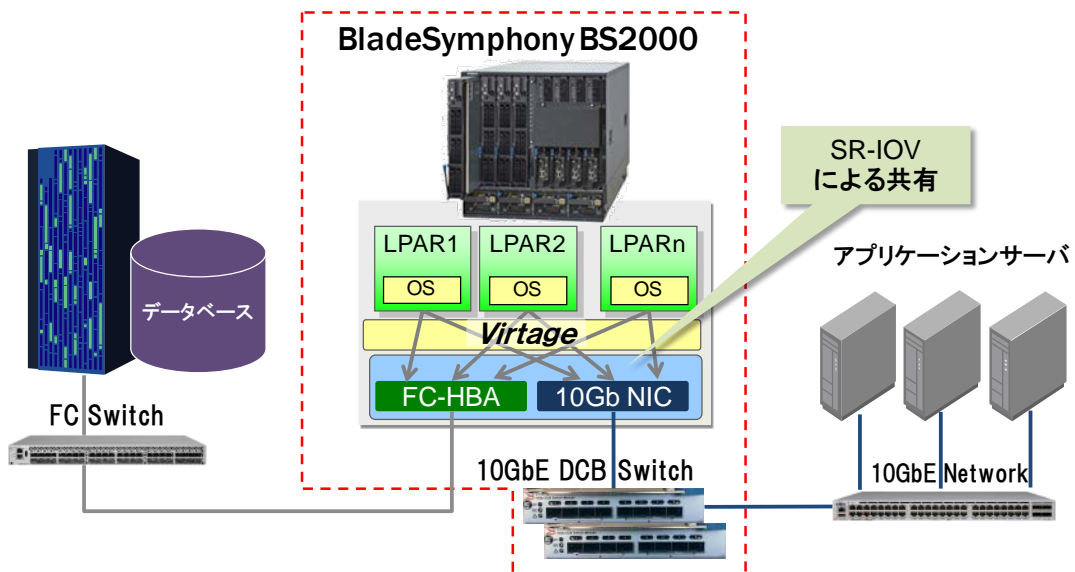
表 3-8. 各方式による共有 NIC 性能比較(ポートあたり)

	ソフトウェアによるNIC共有	SR-IOVによるNIC共有
最大スループット	約 2.2Gbps	約 8~9Gbps
Virtage 内部処理で消費するプロセッサコア数	約 2 コア	0.1 コア以下

4 ユースケース(BS2000 での大規模データベースシステム)

特長

- ① SR-IOV により、各 LPAR あたりの帯域をデバイス占有方式と同等の帯域を確保。DCB スイッチと組み合わせることで、10G ネットワークに対応可能な共有デバイス処理環境と広帯域通信路を確保。各 LPAR とアプリケーションサーバやデータベースとの間で高速かつ低遅延なネットワークを構成します。
- ② Virtage により、仮想 OS から直接ハードウェアにアクセスすることで、I/O アクセス時の性能劣化抑え、高速ストレージや広帯域ネットワークの性能を極限まで活かします。
- ③ SR-IOV 方式のデバイス共有処理によるホスト CPU への負荷は無いので、サーバ(BladeSymphony)上に搭載されるアプリケーションに CPU リソースを集約することが可能です。



5 使用上の注意事項

BS2000 SR-IOV における使用上の注意事項については、下記 URL の「SR-IOV 使用上の注意について」をご参照ください。

BS2000 ユーザーズガイド: http://itdoc.hitachi.co.jp/Pages/document_list/manuals/bladesymphony.html#bs2000

付録 1. ソフトウェアによる NIC 共有

SR-IOV を使用しない場合の NIC 共有方式として、ソフトウェアによる NIC 共有があります。これは複数の LPAR 間で物理 NIC を共有して外部ネットワークにアクセスし、LPAR 間においては物理 NIC を介さずに通信可能としています。

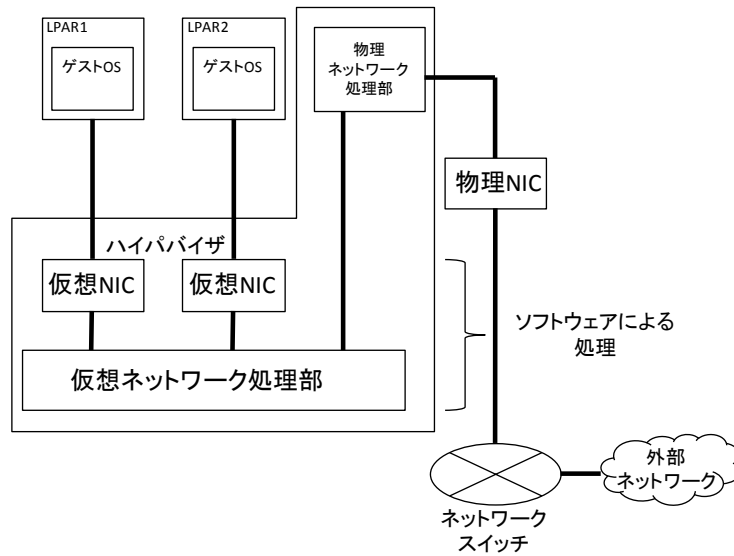


図. ソフトウェア方式による NIC 共有の概略図