

RedHawk と RedHat Linux オペレーティングシステムの リアルタイム・ネットワークパフォーマンスの比較

はじめに：

Concurrent 社の RedHawk Linux オペレーティングシステムは、タイムクリティカルな環境またはハードリアルタイム環境を必要とするリアルタイムアプリケーション向けに設計されています。そのユーザーレベルのコマンドは、Red Hat Enterprise Linux および CentOS と完全に互換性があります。このホワイトペーパーでは、いくつかのネットワーク・インタフェースカード（NIC）上の TCP/IP および UDP/IP プロトコルの平均往復時間を 2 で割った（RTT/2）を調べることにより、これらの Linux カーネルがネットワーク遅延に与える影響を分析します。また、システムに負荷がかかっている場合、Concurrent 社の RedHawk カーネルは、追加の手順やユーザー設定を行うことなく、レイテンシとジッターの点で Red Hat および CentOS カーネルよりも常に優れていることを示します。RedHawk からの追加のユーザーレベルのコマンド（具体的には、shield コマンドと run コマンド）は、Linux カーネル固有の同様のコマンドライン構成メカニズムに匹敵しますが、パフォーマンスが大幅に向上します。



検証環境：

このネットワーク性能試験は、Intel Xeon 3.0 GHz クアッドコアプロセッサを搭載した 2 台の同一のテストマシン上で実行しました。設置されたサブネット上のテストマシン間には、追加のネットワークトラフィックなしで専用のサーバーとクライアントの関係を確立しています。また、テストマシン間の NIC カードを、Mellanox 社パッシブ銅線 InfiniBand QSFP 2m ケーブル（40 Gb/s）を介して直接リンクしました。

このテストは、以下のカーネルと NIC を使用しました。

Kernel：

- RedHawk 7.5.2
- PREEMPT_RT を使用した RHEL 7 RT
- CentOS 7

NIC：

- Solarflare SFN8542 デュアルポート 40Gbit QSFP+ と OpenOnload
- Mellanox ConnectX-5 CX555A VPI/FDR/EDR InfiniBand 100Gbit デュアルポート QSFP28 PCI Express 3.0 x16
- Intel Omni-Path HFI Silicon 100 シリーズ

平均往復時間 (RTT/2) で測定されるレイテンシテストを実行し、TCP/IP および UDP/IP 転送プロトコルを使用して 3 つの異なる NIC の各カーネル間で比較をしました。各カーネルについて、システムに追加の負荷をかけた場合と不可をかけない場合のレイテンシをテストしました。負荷なしのテストは、他のユーザープロセスを実行せずにテストを実行します。レイテンシは、CPU シールドの有無にかかわらずシステムに負荷がかかった状態でさらにテストされます。この実験でのシールドとは、RedHawk カーネル CPU シールド機能の実装 (shield コマンドを使用)、または Linux のコマンドライン制限 `isolcpus` で、テスト用に単一の CPU を分離することを指します。これらの関数は両方ともカーネルによる CPU 使用量をユーザーレベルのコマンドを使用して、特に目的とするタスクに制限することを意図しています。この場合、CPU 1 がシールドされ、すべてのネットワークタスクに割り当てられている間に、両方のテストシステムに負荷がかかります。すべてのネットワークカードには、CPU 1 への割り込みアフィニティも与えられます。単一の CPU ですべてのネットワークタスクを実行するために与えられた特異性は、ベンチマークの目的で行われこの実験でパフォーマンスを最適化する手段としてではありません。

テストマシンは、Solarflare Onload レイテンシベンチマークガイドで推奨されている次のコマンドを使用してベンチマークテストを実施しました：

- `systemctl stop cpupower`
- `systemctl stop irqbalance`
- `systemctl stop firewalld`
- `ethtool -C デバイス rx-usecs 0 適応-rx オフ`
- すべてのネットワークカード割り込みに対して、SMP アフィニティを CPU 1 に設定する
- `tuned-adm` プロファイルのネットワーク遅延
- `sysctl net.core.busy_poll = 50`
- `sysctl net.core.busy_read = 50`

レイテンシデータは、openonload.org にある `sfnettest-1.5.0` の `sfnt-pingpong` ユーティリティを使用して取得されました。本書では、RTT/2 とジッターはナノ秒で報告されますが、マイクロ秒に変換しています。`sfnt-pingpong` ユーティリティは、1 バイトから 65,536 バイトまでの複数のパケットサイズの RTT/2 とジッターを報告します。TCP および UDP 転送プロトコル間でデータをさらに比較しました。カーネル間の比較は、32 バイトのパッケージサイズで標準化しています。

ネットワーク・インタフェースカード：

Solarflare 製 デュアルポート 40Gbit QSFP+ネットワークアダプターは、正確な時刻同期とハードウェアタイムスタンプを 40GB/10GB イーサネットと組み合わせています。また、アダプターは、カーネルバイパスを高速化するユーザーレベルの TCP/IP ネットワークスタックを提供する Onload をサポートし、非常に低いレイテンシを提供します。Onload スタックには、送信バッファと受信バッファ、開いている接続、および関連するポート番号とスタックオブ

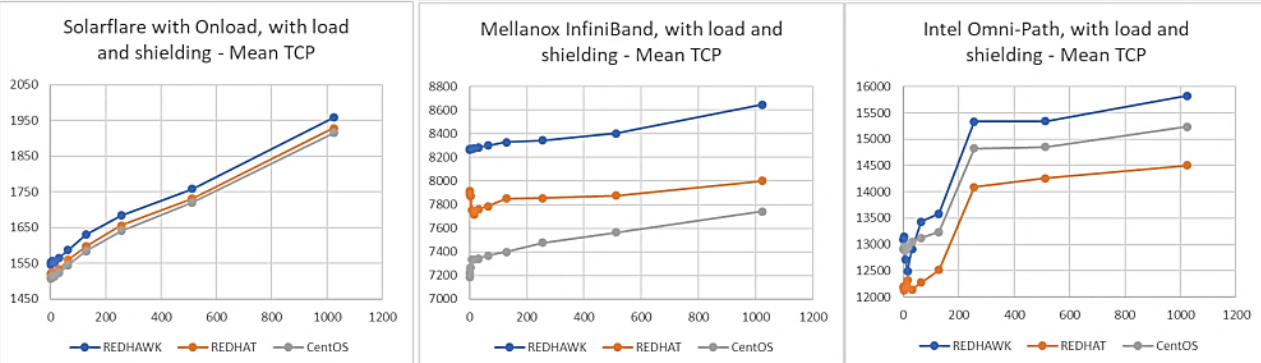
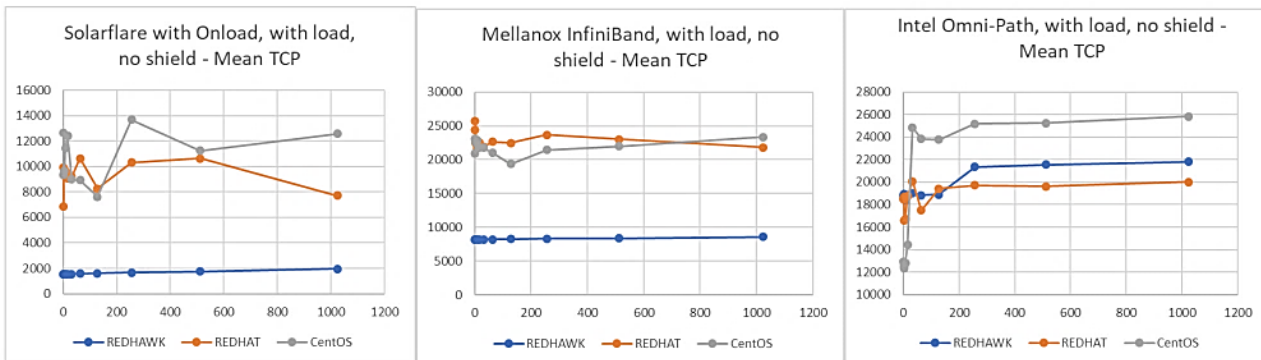
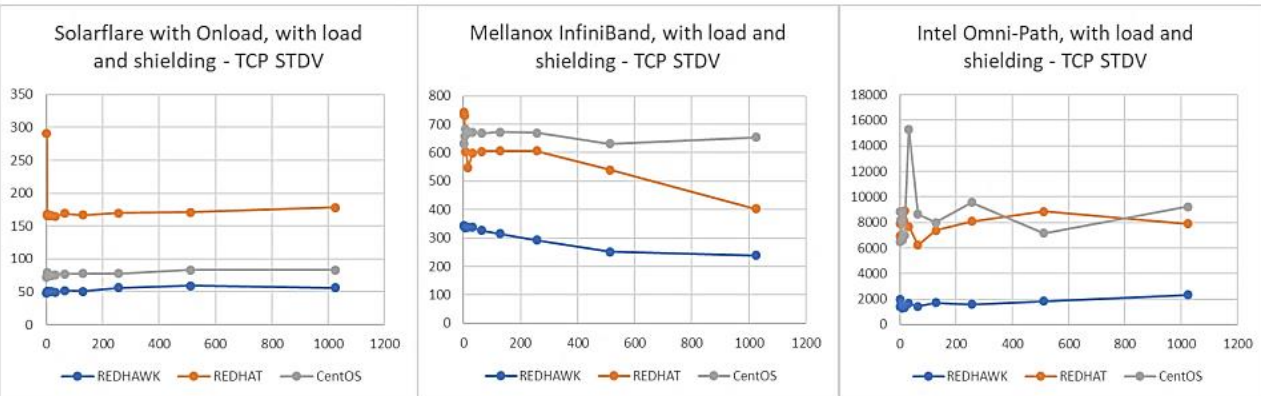
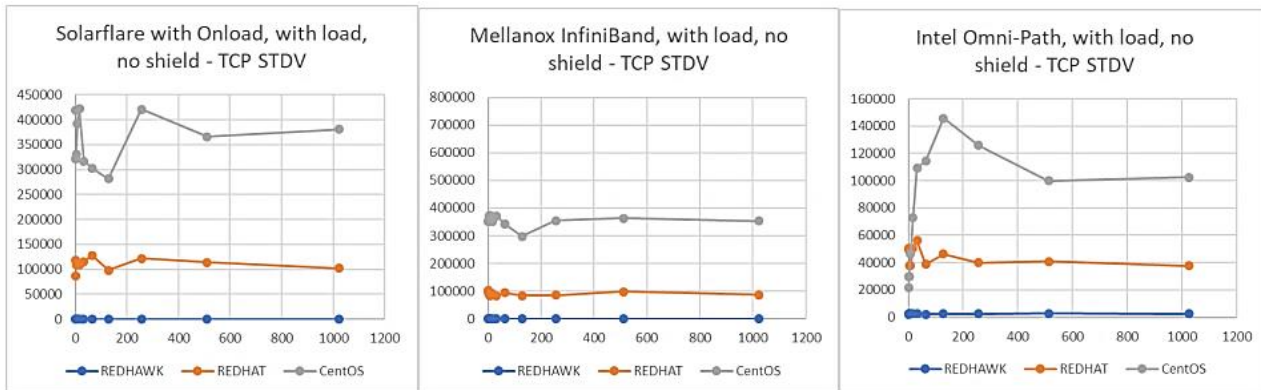
ションが含まれます。Onload は、カーネルからユーザーへの割り込みを排除することで CPU のオーバーヘッドを削減し、ネットワークアダプターがネットワークの受信呼び出しを直接処理する間、アプリケーションをユーザー空間に残し、それらの義務をカーネルから取り除きます。

Mellanox 製 ConnectX-5 イーサネットアダプターカードは、高性能、Web 2.0、クラウド、データ分析、ストレージプラットフォームを最大化するための新しい加速エンジンを促進します。このテクノロジーは、仮想スイッチとルーターのオフロードにより、クラウドと NFV プラットフォームの効率を改善するために転送速度を上げることを目的としています。

Intel Omni-Path ホストファブリックインタフェースシリコン 100 シリーズアダプターは、パケットごとの処理オーバーヘッドを削減し、マルチコアスケーリングを実現する大容量パケット転送サポートにより、効率を向上させます。ホストファブリックインタフェースアダプターは、リンク転送パケット（LTP）のトラフィック制御を緩和して効率的なワイヤー転送を行います。一方、Omni-Path アーキテクチャー（OPA）は、待ち時間とスループットを最適化するスイッチング決定を行うことを目的としています。

各 NIC に対するオペレーティングシステムの比較：

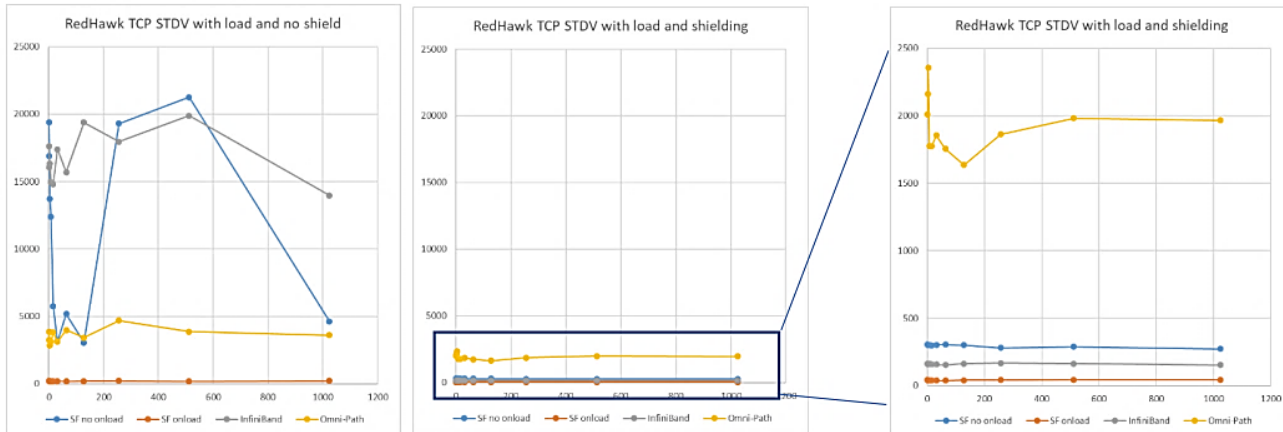
以下に示すデータは、RedHawk のカーネルがシステムに負荷がかかったときに、他のオペレーティングシステムよりも一貫して少ないジッタで相互作用することを示しています。次のグラフは、バイト単位の転送パケットサイズとナノ秒単位の標準偏差時間を示しています。すべてのパケットサイズで RTT/2 倍の競争力を備えているため、ジッターが大幅に減少し、リアルタイムアプリケーションに焦点を当てた一貫性の信頼性が高まります。以下に示すレイテンシ標準偏差の比較は、CPU シールド方法が実装されている場合、RedHawk テストはすべてのパケットサイズで Red Hat および CentOS テストよりもほぼ 50%少ないジッタを一貫して実行していることを示しています。



上のグラフは、パケットサイズをバイト単位で示しレイテンシ RTT/2 時間をナノ秒単位で示しています。システムに負荷があり、シールドがないインスタンスでは、RedHawk は、Solarflare 製および Mellanox 製 InfiniBand アダプターを実装するときに、比較的短い待ち時間を示します。シールドと負荷の両方が存在する場合、3 つのカーネルすべてが

互いに 1 マイクロ秒以内に実行される傾向があります。

オペレーティングシステムに対する NIC の比較 :



RedHawk は、同じ NIC 上の他のオペレーティングシステムと比較して、より優れたレイテンシ標準偏差値を示しますが、Onload を備えた Solarflare 製 NIC (オレンジ色) は他のネットワークアダプターよりもジッターを最も改善します。



NIC	Back-ground Load	CPU Shielding	OS	snft-pingpong (units in microseconds)			
				tcp		udp	
				RTT/2	jitter stdev	RTT/2	jitter stdev
Solarflare without onload	No	N/A	RedHawk 7.5.2	6.997	0.463	6.802	0.234
			RHEL 7 RT	8.452	0.255	7.662	0.242
			CentOS 7	7.014	0.231	6.088	0.207
	Yes	No	RedHawk 7.5.2	7.636	0.29	6.345	0.511
			RHEL 7 RT	114.756	313.876	48.68	195.245
			CentOS 7	21.744	335.4	16.944	284.923
	Yes	Yes	RedHawk 7.5.2	7.844	0.237	6.26	0.465
			RHEL 7 RT	8.475	0.252	7.724	0.321
			CentOS 7	7.122	0.154	6.167	0.135
Solarflare with onload	No	N/A	RedHawk 7.5.2	1.538	0.123	1.431	0.094
			RHEL 7 RT	1.513	0.089	1.416	0.075
			CentOS 7	1.51	0.101	1.412	0.099
	Yes	No	RedHawk 7.5.2	1.545	0.142	1.444	0.118
			RHEL 7 RT	9.219	115.35	3.736	54.921
			CentOS 7	8.982	317.102	3.699	150.493
	Yes	Yes	RedHawk 7.5.2	1.565	0.049	1.439	0.044
			RHEL 7 RT	1.533	0.165	1.427	0.153
			CentOS 7	1.523	0.076	1.423	0.074
Mellanox InfiniBand	No	N/A	RedHawk 7.5.2	8.006	0.219	7.512	0.381
			RHEL 7 RT	8.371	0.879	7.446	0.939
			CentOS 7	7.148	0.656	6.074	0.483
	Yes	No	RedHawk 7.5.2	8.2	0.275	7.686	0.421
			RHEL 7 RT	21.916	84.718	13.975	65.658
			CentOS 7	21.787	371.32	21.427	678.71
	Yes	Yes	RedHawk 7.5.2	8.284	0.338	7.735	0.341
			RHEL 7 RT	7.762	0.599	7.517	0.834
			CentOS 7	7.342	0.672	6.278	0.573
Omni-Path	No	N/A	RedHawk 7.5.2	12.208	0.395	11.341	0.539
			RHEL 7 RT	10.998	9.249	9.637	5.036
			CentOS 7	11.59	0.317	10.193	0.302
	Yes	No	RedHawk 7.5.2	19.028	2.382	16.397	2.973
			RHEL 7 RT	20.081	56.205	12.826	26.895
			CentOS 7	24.837	109.508	19.374	96.61
	Yes	Yes	RedHawk 7.5.2	12.918	1.64	11.779	1.406
			RHEL 7 RT	12.148	7.671	10.404	7.312
			CentOS 7	13.047	15.309	11.26	7.571

Table 1 マイクロ秒単位の snft-pingpong レイテンシ値。NIC、シールド、OS ごとに分類

まとめ：

上記の表は、システムに負荷がかかっているときに RedHawk カーネルが提供するパフォーマンスの利点を強調しています。シールドされた条件下での平均往復遅延時間は同様ですが、RedHawk はさらに、Red Hat リアルタイムカーネルと比較して標準偏差をほぼ 50%削減します。この一貫性の強化は、リアルタイムのアプリケーションパフォーマンスを確保するために必要な信頼性の向上に直接つながります。Onload を使用すると、RTT/2 およびジッターが TCP および UDP プロトコル全体で一貫して向上します。



Concurrent 社について

Concurrent 社は、産業および防衛市場向けの高性能リアルタイム Linux コンピューターシステム、ソリューション、ソフトウェアの業界屈指のプロバイダーです。ハードウェアインザループおよびマンインザループシミュレーション、データ収集、および産業システムに重点を置いており、航空宇宙および防衛、自動車、エネルギー、金融などの業界にサービスを提供しています。コンカレント社はフロリダ州ポンパノビーチにあり、北米、ヨーロッパ、アジアにローカルオフィスがあります。

コンカレント社の RedHawk Linux は、Intel x86 および ARM64 プラットフォーム用のオープンソース Linux オペレーティングシステムの業界標準リアルタイムバージョンです。RedHawk Linux は、タイムクリティカルかつハードなリアルタイム環境で必要な保証されたパフォーマンスを提供します。RedHawk は、モデリング、シミュレーション、データ収集、プロセス制御、イメージングシステムなどの幅広いサーバーおよび組み込みアプリケーション向けに最適化されています。