

10ギガビット・イーサネット環境における低レイテンシソリューション

Arista 7148SX
29West Latency Busters® Messaging(LBM)
Chelsio S310E

パフォーマンス・テスト

2010年6月

ARISTA



発行:

NVC NETWORK VALUE COMPONENTS

エグゼクティブ・サマリ

今日のトレーディング市場では、アルゴリズム取引の急増や市場への直接的アクセス (DMA) の拡大、取引所・PTSの処理能力の向上等により、低レイテンシで大量のメッセージの処理を実現することができるインフラストラクチャが求められてきています。

これは証券会社各社および金融市場に携わる金融機関にとって、取引アプリケーションで競争相手に1マイクロ秒あるいは1ナノ秒でも競合相手より先に行くことが、利益を生む要因となるためです。

本レポートは、10ギガビットEthernet環境において、弊社が考える低レイテンシ・メッセージングソリューションとパフォーマンスについて、説明します。

当該ベンチマークテストでは、Linuxを搭載した汎用サーバ、29Westのメッセージング・ミドルウェア、ChelsioのS310Eアダプタ、Aristaの7148SXスイッチを使用しました。

ベンチマークテストは、16から1024バイトの各メッセージサイズごとに、25,000から2,000,000 Message Volumeを送信し、レイテンシ値を測定しました。

全テスト項目における最小レイテンシ値の平均

21.8 μ sec

全テスト項目毎の平均値

26.0 μ sec

また、Aristaスイッチを介さずに測定した結果は以下のとおりです。

全テスト項目における最小レイテンシ値の平均

20.7 μ sec

全テスト項目毎の平均値

25.1 μ sec

このように、Aristaスイッチを介した場合でも大幅なパフォーマンスの低下がないことを確認しました。

今回はテスト環境の都合上、クライアント・サーバ間の1対1の通信の結果ですが、マルチキャスト通信における1対Nの通信においても同様のパフォーマンスが維持できると予測しております。

テスト結果の詳細

下図のグラフは、メッセージサイズおよびメッセージ数とレイテンシの関係を示したものです。

図1. メッセージサイズごとの最小レイテンシ

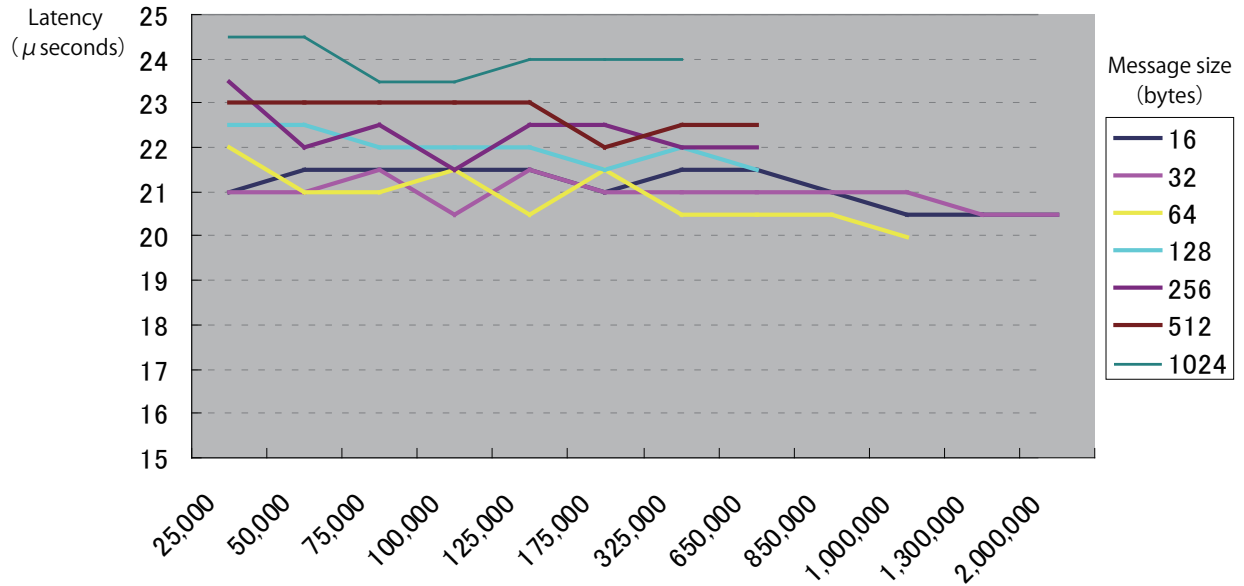


図1より、どのメッセージサイズにおいても、最小レイテンシ値において、メッセージ数による影響を大きく受けずにおよそ2 μsec程度のゆらぎで抑えられていることを計測しています。

図2. メッセージサイズごとの平均レイテンシ

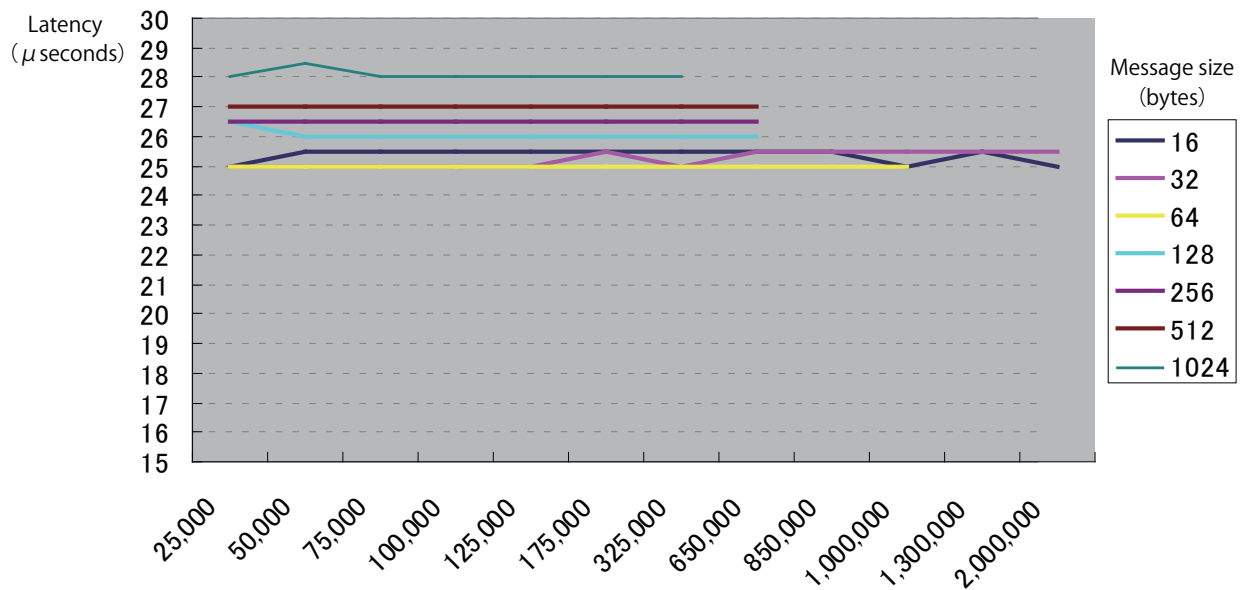


図2より、メッセージ数によらず同様のレイテンシ値を計測しています。これは、大量のメッセージ処理時にも安定して動作することができることを示しています。

テスト項目ごとの結果

| Message Size | Message Volume | Minimum (μsec) | Average (μsec) | Message Size | Message Volume | Minimum (μsec) | Average (μsec) |
|--------------|----------------|----------------|----------------|--------------|----------------|----------------|----------------|
| 16 | 25,000 | 21 | 25 | 128 | 25,000 | 22.5 | 26.5 |
| | 50,000 | 21.5 | 25.5 | | 50,000 | 22.5 | 26 |
| | 75,000 | 21.5 | 25.5 | | 75,000 | 22 | 26 |
| | 100,000 | 21.5 | 25.5 | | 100,000 | 22 | 26 |
| | 125,000 | 21.5 | 25.5 | | 125,000 | 22 | 26 |
| | 175,000 | 21 | 25.5 | | 175,000 | 21.5 | 26 |
| | 325,000 | 21.5 | 25.5 | | 325,000 | 22 | 26 |
| | 650,000 | 21.5 | 25.5 | | 650,000 | 21.5 | 26 |
| | 850,000 | 21 | 25.5 | 256 | 25,000 | 23.5 | 26.5 |
| | 1,000,000 | 20.5 | 25 | | 50,000 | 22 | 26.5 |
| | 1,300,000 | 20.5 | 25.5 | | 75,000 | 22.5 | 26.5 |
| | 2,000,000 | 20.5 | 25 | | 100,000 | 21.5 | 26.5 |
| | | | 125,000 | | 22.5 | 26.5 | |
| 32 | 25,000 | 21 | 25 | 512 | 25,000 | 23 | 27 |
| | 50,000 | 21 | 25 | | 50,000 | 23 | 27 |
| | 75,000 | 21.5 | 25 | | 75,000 | 23 | 27 |
| | 100,000 | 20.5 | 25 | | 100,000 | 23 | 27 |
| | 125,000 | 21.5 | 25 | | 125,000 | 23 | 27 |
| | 175,000 | 21 | 25.5 | 1024 | 175,000 | 22 | 27 |
| | 325,000 | 21 | 25 | | 325,000 | 22.5 | 27 |
| | 650,000 | 21 | 25.5 | | 650,000 | 22.5 | 27 |
| | 850,000 | 21 | 25.5 | | | | |
| | 1,000,000 | 21 | 25.5 | | | | |
| | 1,300,000 | 20.5 | 25.5 | | | | |
| | 2,000,000 | 20.5 | 25.5 | | | | |
| 64 | 25,000 | 22 | 25 | 1024 | 25,000 | 24.5 | 28 |
| | 50,000 | 21 | 25 | | 50,000 | 24.5 | 28.5 |
| | 75,000 | 21 | 25 | | 75,000 | 23.5 | 28 |
| | 100,000 | 21.5 | 25 | | 100,000 | 23.5 | 28 |
| | 125,000 | 20.5 | 25 | | 125,000 | 24 | 28 |
| | 175,000 | 21.5 | 25 | | 175,000 | 24 | 28 |
| | 325,000 | 20.5 | 25 | | 325,000 | 24 | 28 |
| | 650,000 | 20.5 | 25 | | | | |
| | 850,000 | 20.5 | 25 | | | | |
| | 1,000,000 | 20 | 25 | | | | |

結論

これまでアプリケーション間の通信を高速化させるために、さまざまな試みがなされてきました。

29West (Informatica)が以前発表した LBMテスト結果により、ソフトウェアのレイテンシを減少させることは可能であると証明されています。現在では、ストリーミング・メッセージのレイテンシを下げるための要因として、メッセージ配送時にかかるレイテンシをどう減少させていくかが重要な課題となっています。

この課題を解決するひとつの方法として、弊社では低レイテンシのハードウェア機器を利用することに注目しました。

弊社が実施した今回のベンチマークテストでは、AristaとChelsioの低レイテンシ性能および安定した高い処理能力と29westの高いメッセージング処理能力を組み合わせることにより、平均26.0 μ secのパフォーマンスを記録しております。

更なるビジネスニーズに沿ったコンフィギュレーションにより、レイテンシの低下を追及する、あるいはスループットを高めていくことが可能です。

また、これからの10ギガビット・イーサネット環境において、低レイテンシ実現のために今後の利用が見込まれるiWARP (RDMA over TCPを実現する通信プロトコル群) については、別途評価して報告させていただく予定です。

テスト環境詳細

| | |
|-------------------------|--|
| 10GigE Switching | Arista 7148 SX |
| Middleware | Latency Busters® Messaging (LBM) 3.6.2 |
| Linux (Standard Kernel) | CentOS 5.4 |
| NIC | Chelsio S310E |
| Server | Intel Xeon E5504 @2.00GHz |

テスト環境詳細

このシステムでは、2台のワークステーションがArista 7148スイッチを通して接続されています。ワークステーションの構成は以下の通りです。

| Hardware Configuration | |
|------------------------|--------------------------------------|
| Model | VORTECHS VST1D-02 |
| Processors | Intel Xeon E5504 (quad-core) 2.00GHz |
| RAM | 2GB |
| Cache | 4096KB |
| OS | CentOS 5.4 kernel 2.6.18-164.el5 |

ネットワーク仕様は以下の通りです。

| Network Configuration | |
|-----------------------|--------------------------------|
| Switch | Arista 7148 SX, 10GbE |
| 10GbE NICs | Chelsio S310E Firmware 1.4.1.2 |
| Network | 10Gb/s Ethernet |

Arista Networksについて



Arista Networksは2004年に設立され、拠点をCaliforniaのMenlo Parkに置き、Cloud Networkingを提唱し、高性能、信頼性、拡張性、低消費電力を備えた10GbEスイッチ製品を送り出しています。

マーケットをリードするネットワークやサーバ企業から技術者を採用し、豊富なネットワーク技術の経験を元にした技術力で市場をリードしており、金融市場においては、低レイテンシ性能が評価され多くのシステムに導入されています。

詳細につきましては、<http://www.nvc.co.jp/product/arista/> および <http://www.aristanetworks.com/>をご参照ください。

29West (Informatica) について



29West(Informatica)は、金融業界に対してハイパフォーマンス・低レイテンシのメッセージングソリューションを提供するリーディングカンパニーです。

29West (Informatica) が提供するUltra Messaging[®]製品群は、ストリーミング、パースシステム、キャッシュ、キューイングおよびデスクトップ配信の全てをカバーします。従来のメッセージングデザインによるミドルウェア製品に比べて、レイテンシを約10分の1まで低減させ、更にネットワーク帯域の最大限の有効活用を可能とします。業界で最高のパフォーマンスを実現し、詳細なチューニング、制御を可能とするパワフルなメッセージングAPIをモニタリング機能と併せて提供しています。

2004年に製品をリリースして以来、既に140社以上の導入実績を持ち、数多くの証券会社、銀行、ヘッジファンド、プロップファーム、取引所等においてミッションクリティカルなシステムで利用されています。

詳細につきましては、<http://www.29west.co.jp/>をご参照ください。

Chelsio Communicationsについて



Chelsio Communicationsは、堅牢でハイパフォーマンスな実証済みのユニファイド・ワイヤー、ストレージソフトウェア、アプライアンステクノロジーを活用し、ネットワークコンバージェンス、ストレージ、クラスタインターコネクトで業界をリードしています。

スケーラブルで、柔軟なアーキテクチャを持つChelsioの10ギガビットEthernetアダプタカードとマルチ・ポートギガビットEthernetアダプタカードは、ハイパフォーマンスコンピューティングをはじめ様々はアプリケーションに不可欠な優れた低レイテンシと高スループットを提供します。

詳細につきましては、<http://www.chelsio.com/>をご参照ください。

謝辞

本レポートのデータ収集に利用したLBMミドルウェアをご提供いただきました29West社、およびこの性能テストのためにS310E10GbEアダプタをご提供いただきましたChelsio社にお礼を申し上げます。

お問い合わせ

本レポートについてのお問い合わせやソフトウェアおよびハードウェア評価のご依頼は、
弊社オフィス宛にお問い合わせください。

NVC NETWORK VALUE COMPONENTS

<http://www.nvc.co.jp/>

E-mail: sales@nvc.co.jp