

Solace Systems(Solace)の高性能なメッセージングミドルウェア
・アプライアンスの、取引執行におけるローレイテンシー、内部
ディストリビューションにおける高スループット、および外部顧

メッセージングミドルウェア： アプライアンス化の潮流

過去十年間で、FX市場は多くの市場参加者にとって大きな収入源となってきました。BIS(Bank of International Settlements)の試算によると、2010年におけるFX市場の一日の平均粗利益は40兆ドルにも上る見込みとのこと。株式市場と比較すると、株式市場が小さく見えてしまいます。FX取引の多くはグローバルな取引に組み込まれた副次的な効果があるもので、株や債券市場のように投資サイクルに左右されるものではありません。

ブローカーやディーラーは、株やデリバティブ取引において主に金融仲介人となりますが、FX市場では銀行が取引を支配しています。特に多様な通貨や大きな輸出市場が存在するアジア地域の企業や機関投資家は、銀行に対して安定したFX取引をもたらしています。

FX取引を取り巻く環境は、Currenexに代表されるECNやその他の金融機関の電子取引プラットフォームの出現により、市場流動性(リクイディティ)が加速化されると共にその複雑さを増してきました。銀行のトレーディングシステムにおいては、顧客からもたらされる内部的リクイディティへの対応に加え、外部からもたらされるリクイディティへの対応を迫られています。外部FX市場へのアクセスやブライズフィードの集約化、高速処理によるレート配信が、電子取引システム要件の中における重要なポイントとなっています。

また、ヘッジファンドや高頻度で取引を行うHFT(High-Frequency Trading)プレーヤーといった新しいタイプの顧客も市場に参入してきました。彼らは飽和状態にある株式市場に加えて、FX市場も組み込んだマルチアセットでの戦略的取引を行います。これらの新しい顧客からの要求は、従来通りの内容に加え、ハイスピードなFXレートの配信や高スループットでのオーダールーティングを重視しています。

自動取引を行う顧客に対するサービスを提供するにあたり、レイテンシーは最も重要な戦略的技術の考慮点の一つとなっています。機を逃したFXレートの提示は、その提示価格の市場競争

力を奪い去ります。さらには、レート配信や注文処理の遅れに伴い、その注文が不成立とせざるを得なくなるなど、顧客にもネガティブなインパクトを与えてしまう危険性があります。

こうした環境の中、メッセージングミドルウェアは、今までにない戦略的なテクノロジープラットフォームとして位置付けられています。言い換えれば、ミドルウェア・テクノロジーの選択はビジネスアプリケーション戦略に対して大きな影響を与えます。

多様性への挑戦

FXトレーディングシステムにおけるミドルウェアプラットフォームには、幅広い顧客要件への柔軟な対応が至上命題として求められています。様々なソースから得られるリクイディティの情報を瞬時に単一のビューに集約し、他の市場参加者に先んじて取引を行うためには、トレーディングシステムにおける各プロセス間の処理時間を低減することが重要です。その一方で、資産管理やリテールバンキングといった他の部門では、アクセスやシステムインテグレーションの容易性を重視しているケースもあります。

外部顧客のニーズも様々です。中規模の法人顧客は、銀行によって提供されるWebベースのユーザーインターフェイスを使用しますが、大企業の顧客は自己システムとのインテグレーションを行うためのAPI(Application Programming Interface)が必要となるケースがあります。一部の企業のシステムやネットワークが高速なFXレートの配信に対応できない可能性があるのに対し、HFT(高頻度取引: High Frequency Trading)に従事する顧客ではこれらの機能を強く求めているケースがあります。

多くの顧客に対してFXレートを一斉配信することは、技術面においての大きなチャレンジです。複雑な組み合わせによるクロスレートや、期日、行使価格などの諸々の要素によるデータ量の増加は、内部及び外部のシステムに限りあるキャパシティの脅威となります。加えて、特定の顧客グループに対してマージンプライスを流すことにより、リアルタイムのデータレートは

更に増加します。高度で先進的なトレーディング・プラットフォームにおいては、数十万通りもの複合的FXレートの配信が求められるケースもあります。それとは対照的に、一般的な株取引で求められる相場情報配信では、FXと比較して少ない情報をより多くの顧客に対して配信することが求められます。株式で取引される株式の種類はある程度限られており、また全ての顧客に対して同じ価格情報を配信することが求められます。

これらの条件を全て満たすFX配信プラットフォームを構築・実装するのは、とても大きなチャレンジであると言えます。

株取引に関して言えば、一般的には以下の3つの要件に対応した複数のメッセージングを利用しています。

- 1) パフォーマンスに敏感なDMA(Direct Market Access)
- 2) 一般的な取引システム間の連携
- 3) 株価・市況情報の配信

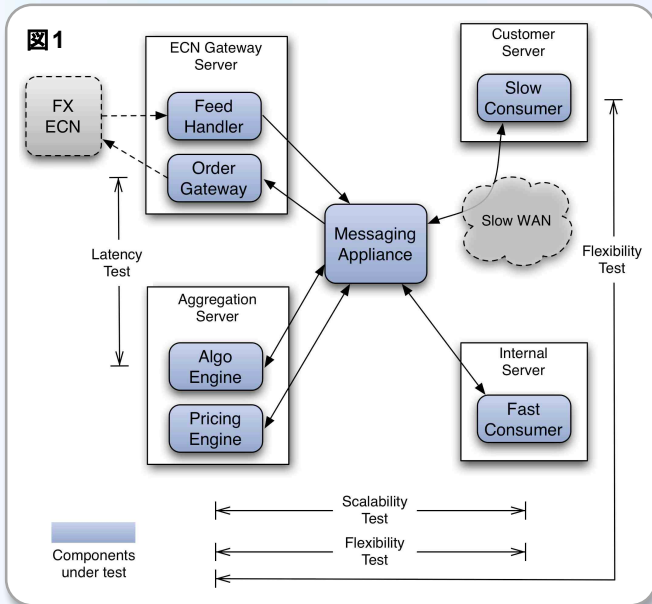
こうした複数プラットフォームのデザインでは、構築・運用・管理におけるコストの増加が懸念されます。コスト削減やデザインの複雑性を最小限に抑えるためには、全てのテクニカルな条件を満たす単体のミドルウェア配信プラットフォームを使用することが理想的とされてきました。幸いなことに、現在ではハードウェア技術を活用したメッセージング・プラットフォームがこうしたビジョンを現実のものに近づけています。

次世代型ミドルウェア

過去20年間において、一般的なトレーディングシステムのデザインは、信頼性のあるブロードキャスト・プロトコルをベースとしたソフトウェアベースのメッセージングミドルウェアを使用し、アプリケーション間の連携を実現してきました。信頼性のあるブロードキャスト型データ配信は、元来UDP(User Datagram Protocol)と呼ばれるコネクション型のPoint-to-Pointデータ配信プロトコルから発展したものでした。一方のTCPをソフトウェアベースで実装したPoint-to-Pointデリバリーでは効率的なスケールアウトが出来ず、最後のメッセージ受信が最初の受信に比べて大きく遅れるなどの不均等な配信という問題を引き起こしていました。

一つの送信者(Publisher)から複数の受信者(Subscriber/Client)への配信を行う一般的なマーケットデータ配信においては、UDPブロードキャストは効率的な実装でした。しかし、UDPでは全てのデータをブロードキャストするため、受信者側で不要なメッセージのフィルタ

FX取引は、アジアの金融機関に大きな収益をもたらしてきました。そしてその収益の増加は、より多くの市場参入者とトレーディング・テクノロジーの高度化を推し進めてきました。そうした中、ハイパフォーマンスなメッセージングミドルウェア基盤の整備は、市場競争力を維持するため最重要課題のひとつとして注目されています。



リングを行うための負担が掛かっていました。UDP/Multicastによる配信ではこの点に改善がなされ、データ配信グループ(Multicast Group)をより細かく区分することにより、部分的ながら先の欠点の抑制を図りました。

この問題はFX取引に大きく関係してきます。FXではそれぞれの顧客に対して異なるプライスが生成され、それぞれのアプリケーションはそのタイミングで利用可能なデータにアクセスする可能性があります。同じ情報を受け取る顧客数が限定されているため、Multicastで対応した場合には非常に多くのMulticast Groupを管理することになります。また、Multicast配信は柔軟性に乏しい側面があり、データ配信に先立って事前に特定のマルチキャストグループを割り当てる必要があります。そして数十・数百・数千にも及ぶ大量の固有のデータ配信を任意のMulticast Groupに振り分けるのは、容易な作業ではありませんでした。

信頼性のあるブロードキャスト・プロトコルの導入に対して、ネットワーク機器の性能面・機能面での進化により徐々に解消されはじめ、その結果としてUDPベースの配信プロトコルの発達へと繋がりました。例として、スイッチドネットワークではそれぞれのサーバーのエンドポイントにて専用の帯域幅を提供し、特定のサーバー向けのトラフィックが同一セグメント上の他のサーバーに与える影響の極小化を実現しました。同様のアプローチにより、ミドルウェアもハードウェア・アプライアンス化されることにより、従来のソフトウェアベースのミドルウェアが抱えていた多くの課題を克服しました。FPGA(Field Programmable Gate Array)やNP(Network Processor)の技術を利用するこれらのアプライアンスは、サーバーというよりネットワーク・ルーターのように位置付けられています。フィルタリングや配信の機能をソフトウェアではなくシリコンに実装することにより、集中型配信サーバー構築に伴う性能面や機能面での制限を取り除くことに成功しています。こうした進化により、改善されたネットワークとミドルウェア・アプライアンス技術との融合により、旧来のUDP Multicast型配信に代わってTCPベースのPoint-to-Point型配信が効率的でより実用性のある代替手段となったのです。

あらゆる要件に対応するマルチパスなミドルウェア・アプライアンスは、従来のサーバーベースのミドルウェアに対して、多くのアドバンテージをもたらしました。ハードウェア処理により、レイテンシーの揺らぎ幅・スパイクを大幅に減少させました。また、到達性保障型のメッセージングにおいても、専用電源で補助された不揮発性のメモリーを使用することにより、従来の汎用サーバー上で稼動するソフトウェアよりも大幅にローレイテンシー化、高スループット化しています。更に、それぞれのエンドポイントでのメッセージの流れは、一元的にモニターして管理されます。そしてメッセージの速度はそれぞれ独立して管理されるため、超高速サーバーやWANを跨いだ先からアクセスする「Slow Consumer」が混在したとしても相互影響はありません。

しかしながら、これらすべての利点を持ってしても、ミドルウェアのアプライアンス化の真の効果はその実装に依存します。集中型のディストリビューションのアーキテクチャーは、ルーティングやフィルタ

リングのような機能性と拡張性の実装を経て初めて実用的となり得ます。

これらの点を念頭に置き、様々なベースにおけるミドルウェア・アプライアンスの稼動についての検証を行いました。マイクロ秒単位のレイテンシーと、数千件/秒を超えるメッセージレートにおいても、Solace Systemsの3260メッセージ・ルーターのパフォーマンスは理想的なものでした。金融業界において急速に受け入れてきたSolaceのメッセージング・アプライアンスを用いて、集中型配信アーキテクチャーが、FXトレーディングアプリケーションで必要とされるパフォーマンスや柔軟性の提供が可能かを考察するテストを行いました。

ソフトウェア構造

図1にメッセージング・アプライアンスを基にしたFX取引の概念構造図が示されています。1つ、もしくは複数のFX-ECN（電子証券取引ネットワーク）が、レート配信フィードとオーダー・ゲートウェイのインターフェースを提供しています。FXトレーディングシステムは、これらのフィード・ハンドラーやオーダー・ゲートウェイを介してECNや他のトレーディングシステムとの通信を行います。アルゴ・エンジンは複数の流動性資産のプライスを総計し、可能な流動性資産に対して自動的に戦略的取引を行います。一般的に、フィード・ハンドラーやオーダー・ゲートウェイは、Active/Passiveなハイアベイラビリティ構成を組んでいます。シンプル化を図るため、図1ではこれらの冗長構成については割愛しました。

プライシング・エンジンは原(Raw)レートまたは集約(Aggregated)レートを受け、補間計算の上で、広範囲に及ぶ先日付の統合的なクロスレートを生成します。またオプションにおいては、複数の行使プライスや行使期限ごとのオプションプライスを生成します。一般的にプライシング

図2: レイテンシーテスト

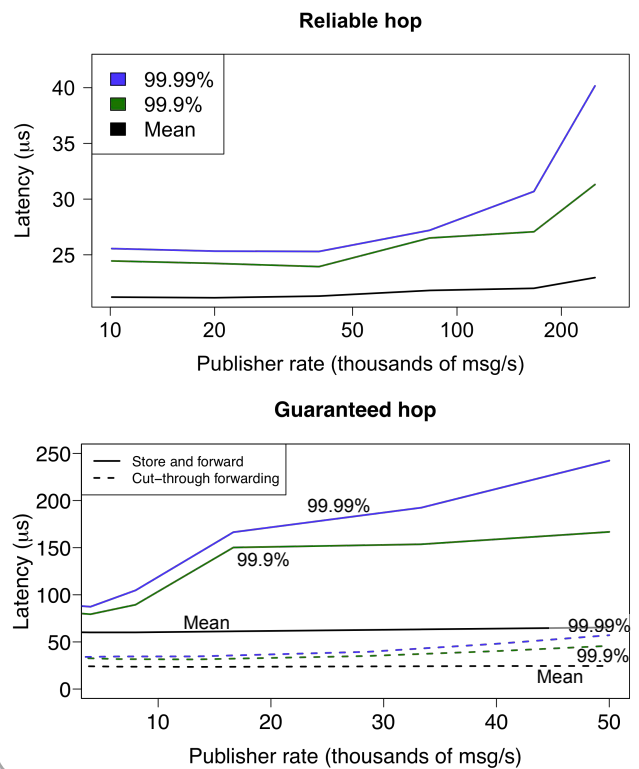
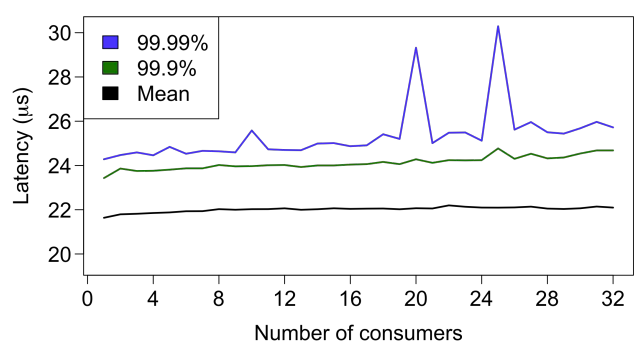


図3: 拡張性テスト



シナリオ	目的	データ送信元 (Producer)	データ受信先 (Consumer)
レイテンシーテスト	-プラットフォームのレイテンシーを測定 -フィード・ハンドラーから配信型(Reliable Messaging)でレート配信 +アルゴエンジンでレートを受信して処理 +到達性保障型(Guaranteed Messaging)で注文を送信 +オーダー・ゲートウェイで受信	-1xフィード・ハンドラー	-1xアルゴエンジン： 5つのレート情報を受信する度に注文を生成
拡張性テスト	-配信型(Reliable Messaging)で大量のレートを配信 -フィルターによるアクセス制御をした場合も計測	-1xプライシング・エンジン	-32のクライアントプロセスでそれぞれ異なるレートを受信
柔軟性テスト	-配信型(Reliable Messaging) -処理能力の高いサーバーと低いサーバーの混在環境での性能を計測	-1xプライシング・エンジン	-1x高速処理サーバー、 -1x低速処理サーバー、 -1x低速処理サーバー：Solaceにて“間引き”を設定

・エンジンはActive/Active型の負荷分散構成を組み、拡張性と冗長性を確保しています。また、インタラクティブなプライシング・エンジン実装の場合には、クォート要求を受けてレートを配信するケースもあります。別のプライシング・エンジン実装では、原レートの更新情報受信の都度レート計算を実行し、継続的にレートの配信を行うようデザインされているケースもあります。配信されたレートは概ね一秒以内の取引有効期限が設定されています。

アプリケーションデザインの観点から見ると、数多くのレートを自動的に生成し配信を行うアプローチは合理的で理解しやすいものです。しかし、その実現には多くの顧客に対して個々に異なるレートを配信するために、高度なルーティングやフィルタリング機能を備えた強固なメッセージング・プラットフォームが必要不可欠です。ある顧客はできる限り早くレートを受信し、高頻度でオーダーを出す「Fast Consumer」である場合があります。また別のクライアントアプリケーションはシステム容量的に高頻度のデータ処理に対応ができないケースや、あるいはWAN(Wide Area Network)を介してデータの交換を行うために処理可能なデータ量に限界のあるような「Slow Consumer」である場合もあります。多くのクライアントアプリケーションは、この極端な2極間のどこのポイントに分類されます。

テスト・シナリオ

このテストでは、FXトレーディング・プラットフォームにおける主要な問題点を調査・計測することを目的としました。

- 1) メッセージングインフラストラクチャーにおけるレイテンシーとスループットの特徴
- 2) クライアント数増加に伴うパフォーマンスへの影響
大量トピック(=レート)の配信能力
- 3) 「Slow Consumer」と「Fast Consumer」が混在する場合の影響

テスト環境の論理構成と各計測ポイントは図1に示してあります。灰色で示されている要素はテストに含まれません。以下の文中にて称される、フィード・ハンドラー、オーダー・ゲートウェイ、アルゴ・エンジン、そしてプライシング・エンジンは、これらの機能を指します。

全てのテストにおいて、テストメッセージがテストコンポーネントを通過する際に、TS-Associates社の「Application Tap card」と「TipOff」の監視アプライアンスを使用し、タイムスタンプをしました。このコンフィギュレーションは、テスト測定結果の影響を最小限にしなが、論理クロックの同期問題を回避します。テストアプローチに関するさらなる詳細については、以下のURLで公開しています。[<http://www.catena-technologies.com/reports/hp-test-approach.pdf>]

テスト結果はグラフをご覧ください。

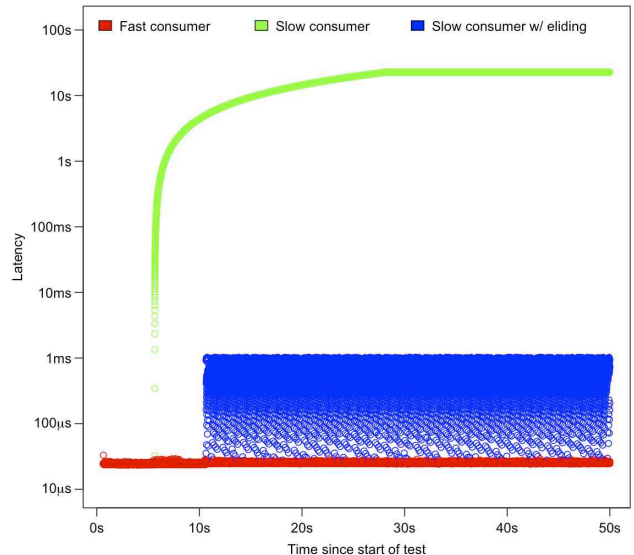
レイテンシーの平均値は、テスト中に受信した全てのメッセージのレイテンシー分布に基づいて平均計算されています。「99.9%(パーセンタイル)」の計測結果は、テスト中にレイテンシーが計測された全メッセージのうち、0.1%のメッセージを除いた場合の最大レイテンシーの値です。(例：1,000件のテストメッセージのうち、最大レイテンシーを記録した1件を除いた999件の中での最大レイテンシー値)同様に「99.99%(パーセンタイル)」の計測結果では、わずか0.01%のメッセージを除外した場合の最大レイテンシーの値を示しています。このアプローチにより、様々なテスト環境において、極端なレイテンシーの値がどのように分布するかを確認しま

した。最初に、リクイディティ・プロバイダーから受け取ったリアルタイムのレートを基に、自動的にオーダーが生成されるまでをシミュレートした「レイテンシーテスト」を行いました。フィード・ハンドラーにより生成されたレートは、信頼性の高いTCP/IPベースの配信型：Reliable Messagingを用いて、アルゴ・エンジンへと送信されます。アルゴ・エンジンは5つのレート更新を受信する毎にオーダーを生成し、到達性保障型：Guaranteed Messagingを用いてオーダー・ゲートウェイへと送信します。その各々のステップ毎にレイテンシーを計測しました。このテストでは基本的なメッセージング性能を計測すると共に、取引前のオーダー管理という用途において到達性保障型：Guaranteed Messagingが十分なパフォーマンスを提供できることを証明することを目的としています。また、ここではStore & ForwardとCut-Throughという2種類のGuaranteed Messagingのモードがテストされています。

2つ目のテストは、プライシング・エンジンから数多くの顧客に対して多種多様のトピック(=レート)を大量に同報配信した場合のパフォーマンスを計測する「拡張性テスト」を行いました。このテストでは、多様且つ大量な配信が求められる環境において、メッセージング・アプライアンスがハイボリュームなデータの配信と並行して効率的なフィルタリングを行うことができる能力を検証しました。最大32のクライアントに対し、各々が固有に500トピック(500種のレート)のセットを取り込み、合計でそれぞれ2,000件/秒のメッセージを受信します。(例：トピックごとに4件/秒のメッセージを受信)

3つ目のテストは、到達性保障型：Guaranteed Messagingにおいてメッセージング・アプライアンスが「Fast Consumer」と「Slow Consumer」が混在する環境でも効率的に能力を発揮できることを確認する「柔軟性テスト」です。基本的な構成は「拡張性テスト」と同様ですが、このテストでは以下3つのコンシューマー・プロセスを稼働させました。

図4: 柔軟性テスト



テストの結果、Reliable Messagingでは平均22~25μ秒、Guaranteed MessagingでもCut-throughモードでは平均25μ秒という結果がでました。また、クライアント数やトピック数の増加が与えるシステムへの影響はほとんど見られませんでした。

- 1) 毎秒500件受信する「Fast Consumer」
- 2) 高頻度のアップデートメッセージを受取る容量のない「Slow Consumer」
- 3) “間引き制御”を受ける「Slow Consumer」

“間引き制御”では、メッセージング・アプライアンス上にて、任意のクライアントに対するメッセージの流量を制限する機能です。もし制限値を超える頻度でデータが配信される場合、メッセージング・アプライアンスは任意の設定時間枠の中の最新のメッセージのみを配信し、その前のメッセージを間引きます。

テスト結果

レイテンシーテストの結果が図2に示されています。フィードハンドラーからアルゴエンジンまでのReliable Messagingの遅延は、毎秒250,000件のレートまで負荷をかけた状態でも23μ秒を超えることはありませんでした。

(注：毎秒250,000件は、使用サーバ機材におけるシングルスレッドのアプリケーションにおける理論上の性能限界値です)

99.99パーセントの計測結果では、依然として低い値ながら相応の増加が確認できます。

一方、アルゴエンジンからオーダーゲートウェイまでの注文情報回送を想定したGuaranteed Messagingでは、高速なCut-throughモードを使用した場合には上述のReliable Messagingとほぼ同等の24μ秒という結果が出ています。平均値は25μ秒以下で安定的に維持し、99.99パーセントの値では相応の増加が見られています。

また、従来のStore & Forwardモードを使用した場合、ルータ上での同期的永続処理に伴い、平均値で66μ秒という結果が出ています。

図3にある拡張性テストの結果からわかる通り、ある程度の負荷をかけた状態でサブスクライバ・トピックやコンシューマー数の増加を行う

ことは、レイテンシーには大きな影響を与えることはありませんでした。1つのコンシューマーで500トピックを登録、合計で2,000件/秒のメッセージを受取った場合の平均レイテンシーは21.6μ秒でした。また、そのコンシューマー数を32として、トピック数や更新レートも比例して増加させた場合、平均レイテンシーは22.1μ秒となりました。グラフ上、99.99%(パーセントタイトル)レベルではややばらつきが見られましたが、それでもいずれのケースにおいて30.3μ秒を上回ることはありませんでした。それら計測数値の揺らぎに関しても、あくまでも瞬間的な揺らぎでした。なお、今回のテストは2台の物理サーバ上で行ったため、CPUの性能限界も踏まえて最大コンシューマー数を32に設定しました。

図4は柔軟性テストの結果を示しています。これは他のグラフとは異なります。合計ではなくそれぞれ個々のレイテンシー測定結果が描かれています。「Fast Consumer」向けのレイテンシー分布を示す赤いポイントは、一貫して低い値で安定しています。「Slow Consumer」向けの測定結果を示す緑色のポイントは、急激な上昇の後にも継続してレイテンシーが増加し、最終的にはメッセージング・アプライアンス上の送信バッファが溢れて滞留メッセージの破棄が発生するまで増加傾向を継続していることを示しています。

また、“間引き処理”の施された「Slow Consumer」は青いポイントで示されています。今回の間引き設定では各トピックの最大更新頻度を「1,000件/秒」と設定しているため、各トピックのレイテンシー値は決して1ミリ秒を大きく超えることはありませんでした。

全てのテストにおいて注目したい点ですが、メッセージング・アプライアンスの性能限界に到達する前に、用意した2台のテスト・サーバ側の性能限界に達してしまいました。

結論

FXトレーディング環境において、ハイパフォーマンス、高い拡張性、そして安定したレイテンシーというのは必要不可欠な要素です。テスト結果が示すように、Solaceのメッセージング・アプライアンスはこれらの資質をしっかりと兼ね備えています。例え「Slow Consumer」が存在した場合においても、レイテンシーの値を予測の範囲内に維持することができます。また、現在のところ到達性保障型：Guaranteed Messagingは配信型：Reliable Messagingよりも高いレイテンシーを示していましたが、Solaceでは2011年中に“カット・スルー型”の到達性保障メカニズムを実装する予定であり、さらに高速なGuaranteed Messagingを実現する見込みです。

Reliable MessagingやGuaranteed Messagingに加え、水平的スケールアウトを実現する拡張性や、WAN(Wide Area Network)への効率的な配信、ハードウェアベースで実装された冗長化機能群、そして「Web Messaging」と呼ばれるWebストリーミングなど、Solaceのメッセージング・アプライアンスはその他数多くの機能をサポートしています。これらの機能群はFX取引やその他の金融サービス・アプリケーションにとって非常に魅力的なものであります。その適応性を証明するために、改めて包括的なテストを行う必要があります。これからも、我々はアプライアンス化の潮流がどのような進化を見せるのかを、見守りたいと思います。

Catena Technologiesは、このテストを実施するにあたりTipOff®とApplication Tap®のハードウェアをご提供いただきましたIS-Associates社のご協力に感謝いたします。

このレポートはSolace Systems, Inc.と合同で準備いたしました。このドキュメントの全著作権は各オーナーに帰します。なお、日本語訳にはセリスソリューションズ社のご協力を頂きました。

Catena Technologiesはシンガポールに本社を持つフィナンシャル・テクノロジー・コンサルティング企業であり、ハイパフォーマンスコンピューティングをベースとしたシステムの構築、デザイン、そしてサービス実装をサポートしています。

OVERALL RESULTS

Test	Configuration	Max. publishing rate	Average latency	99.9% latency	99.99% latency
Latency test	Reliable hop	250,000 msg/s	23.0 μs	31.3 μs	40.2 μs
Latency test	Guaranteed - store & forward	50,000 msg/s	65.4 μs	166.8 μs	242.3 μs
Latency test	Guaranteed - cut-through	50,000 msg/s	24.5 μs	45.9 μs	57.1 μs
Scalability test	32 consumers	64,000 msg/s	22.1 μs	24.7 μs	25.7 μs
Flexibility test	Fast consumer	1,000 msg/s	25.2 μs	27.4 μs	28.6 μs
Flexibility test	Eliding consumer	1,000 msg/s	526 μs	1,026 μs	1,028 μs
Flexibility test	Slow consumer	1,000 msg/s	16,600 ms	22,600 ms	22,700 ms

TECHNICAL SPECIFICATIONS

Messaging Appliance		Test Servers	
Model	3260M Message Router	CPU	2 x Intel® Xeon® X5550 2.67GHz 4 cores per socket
Operating system	SolOS-TR Version 5.2.1.3	Memory	12 GB
Test Network		Operating system	CentOS 5.2
Network switch	Arista 7148SX	Benchmark software	Custom developed C test applications
TipOff® appliance	TipOff® Stack 2.0.7 Application Tap® firmware 1.1.8; driver 1.0.6	Network interface	Solarflare SFN5122
		Kernel bypass	OpenOnload 20100604-u2