Unified Messaging Infrastructure for Global FX Trading Platforms

FX trading platforms have some unique requirements driven by the continuous (24x6) and globally distributed nature of trading, tiered pricing and massive amounts of money involved.

This paper outlines the typical functions and data flows of an FX trading platform, which includes the non-persistent delivery of pricing and persistent delivery of trades to internal and external recipients over LANs, WANs to desktop and mobile applications.

Today investment banks use different messaging and infrastructure technologies for each of these flows because until now no single technology has been able to meet all their needs. Not only is the resulting disjointed technology fabric expensive, complicated, unstable and hard to manage, it makes it hard to develop and deploy new features, and very difficult to maintain consistent delivery latencies.

This paper explains how Solace technology addresses the varied requirements of all flows in a global FX trading platform in a best-of-breed manner while offering a more manageable, robust and secure environment with built-in high availability and disaster recovery functionality.

 http://www.solace.com

Introduction

Solace has advanced the state of the art of messaging middleware by introducing purpose-built message routers that can distribute massive amounts of information over all kinds of networks and using a variety of qualities of service/persistence. As such Solace's hardware is uniquely capable of improving the performance, reliability and manageability of FX trading platforms while simultaneously reducing expenses and enabling greater flexibility.

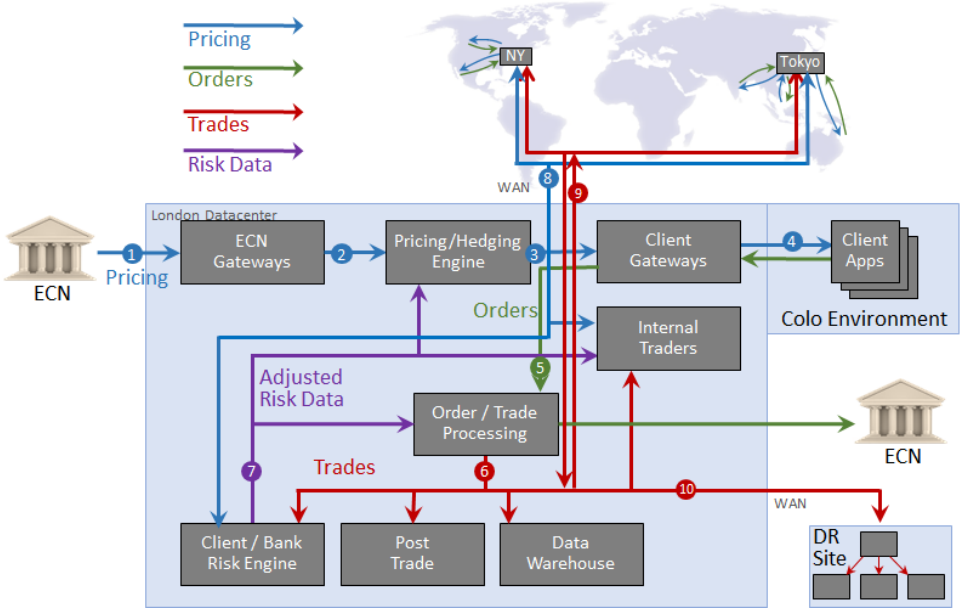
- **Performance:** Solace offers very high performance non-persistent and persistent messaging, both of which are used in FX platforms for various flows of pricing, order, risk and trade data. By upgrading cards inside the Solace message router, vertical scalability can be achieved where performance can be doubled (or more) by replacing one card with a faster version of that card – compared to horizontal scaling of brokers, which is the only option with software.
- **Robustness and Stability:** Solace's hardware broker-based architecture provides true decoupling between publishers and subscribers, ensuring that slow, offline or misbehaving consumers never affect the stability, predictability or manageability of the overall system.
- **Simplicity:** Solace supports all messaging qualities of service and efficiently distributes data over local and wide area networks so you can manage the distribution of pricing, risk, order and trade data within your datacenters and colo environments, over oceans and directly to traders, all with a single platform. This results in a simpler architecture that is more predictable and robust than when many technologies are cobbled together.
- **More Efficient Application Development:** With Solace technology, each application uses a single API to access any type of data in the system (pricing, orders, trades, risk) with any quality of service (persistent/non-persistent) produced anywhere in the global platform – so application developers need not deal with separate APIs and messaging system behaviors. Similarly, the sophisticated features built into Solace message routers ensure that developers spend their effort on applications – not infrastructure.
- **Rich Real-Time Management:** Solace message routers provide detailed, integrated real-time status and statistics from layer 2 to layer 7 that cannot be produced by software systems and allow our clients to use the message router as the central monitoring point for application health. This rich information is also used for troubleshooting and capacity planning of the system and provides unmatched operational transparency.
- **Integrated Functions:** Solace's platform integrates functions such as high availability, replication for disaster recovery without storage replication infrastructure, WAN optimization, and adaptive rate control delivery – all providing a simpler environment with fewer moving pieces under the administrative control of the platform owner instead of other service teams.

No other product can offer all of this functionality – software brokers can't maintain high performance and robustness in volatile high-volume real-world scenarios, and peer-to-peer solutions don't offer the manageability, stability, simplicity, or security it takes to power global FX trading platforms.

Functions and Data Flows in a Global FX Trading Platform

The details of FX trading platform architectures vary from firm to firm, but the general functions and flows are very similar across architectures. This section describes these common components, the types and requirements of the flows between them and how Solace messaging message routers meet these needs.

The figure to the right shows the common functions and flows of a global FX trading platform, which can generally be classified as follows:



Price Flows

The distribution of pricing information (1) includes the receipt of normalized ECN pricing from ECN gateways (2) and subsequent distribution of internal house (best) pricing for currency pairs produced by pricing engines (3), and sometimes client-specific or tiered pricing which can be generated by pricing engines themselves or by the client gateways. This price distribution over the LAN, often in a colocation site, sees a moderate peak rate of hundreds of thousands of messages a second, and demands the ultra-low latency of non-persistent messaging.

Pricing data (typically ECN or house prices) is also typically distributed over the WAN (8) between trading sites such as New York, London and at least one site in Asia like Hong Kong, Tokyo or Singapore. These WANs have much less bandwidth than LAN links, feature high round trip times (around 100ms NY to Europe, 180ms NY to Asia), and suffer from much higher packet loss ratios than LANs, but must still satisfy the requirement to deliver the data with the lowest possible latency since the value of pricing is time sensitive. As a result, these pricing flows benefit considerably from WAN optimization techniques and typically use TCP due to the many problems associated with using reliable multicast over the WAN. Pricing to customers (4) is usually sent from client gateways using FIX or a proprietary protocol.

Pricing information is also often distributed to desktop applications used by internal traders. Their needs are different than those of server applications – desktop applications need to receive pricing at a slower rate in order to not overwhelm the computer or network and also because humans, and even Excel, can't handle the same update rate as servers. Using multicast for these flows is problematic because the networks to the desktops are less well controlled/engineered, there is typically little overlap in interest and the desktops themselves being less controlled can often become slow consumers. Using multicast to send data over such networks doesn't satisfy security audit requirements. Caching is usually required so when traders first subscribe to a currency spread they can receive the last value of the spread, or get a detailed order book of various prices.

Order Flows

Clients submit orders via client gateways using FIX or a proprietary protocol. There are many variations in how these are handled by the platform – some firms execute trades within the client gateways while others forwarded orders (5) to an order/trade processing function as shown here. If forwarded between platform components as part of front office processing, orders require low latency delivery but also often require persistence to ensure they are not lost in transit. These orders typically stay within a single trading location and are not sent over the WAN to other trading venues.

Sometimes client orders are also stored in data warehouses or processed in real-time by functions other than the trading platform, introducing the need for a reliable pub/sub infrastructure. Internal traders also need to be able to inject or adjust orders in the system. Depending on the architecture, either persistent or non-persistent (via request/reply) are used for these flows.

Trade Flows

As trades occur, trade events are emitted into the platform (6) for distribution to many local and remote functions. As part of the front office function, trade distribution needs to be handled with ultra-low, consistent latency between local components, and demands persistence as this completed trade cannot be lost for any reason.

As part of post trade processing, these events also require persistent messaging to ensure lossless delivery and fanout to many post-trade processing functions. Fanout could be to applications within the colo or remotely in the bank's datacenter. Handling of slow consumers is critical so slow functions like trade data warehouse don't impact real-time consumers such as hedging and risk/spread adjustment processing which require low, consistent latency. Support for propagation to a disaster recovery site (10) is typically required, as well as propagation to other sites over the WAN (9) for functions like risk calculation.

Post-trade systems are often located on the bank's premises and sometimes centralized in one geography. Here again you need a distribution infrastructure that provides WAN optimized, secure and lossless delivery.

Finally, as new requirements are constantly being mandated due to regulatory and internal risk initiatives, it is important to have a robust, flexible, high performance trade distribution fabric so that effort can be spent satisfying new requirements instead of dealing with the data movement infrastructure.

Traders often need to receive trade completion events – ether for orders they initiated or to monitor the progress of some flow in the system. These trade notifications can be delivered as either persistent or non-persistent messages depending on the architecture of the application.

Risk Data Distribution

As trades occur, updated risk positions are computed and are published to the platform (7) in order to influence subsequent pricing and trading activities. These flows are typically non-persistent to applications in the same trading venue – so they are like pricing flows but are often larger and less frequent than price updates.

Applying Solace Technology to Global FX Trading

Solace provides a single technology to address all the data distribution needs of Global FX trading platforms discussed in the previous section. Using a single infrastructure for all these flows provides significant ongoing operational savings by having a single platform for management, monitoring, capacity planning, troubleshooting and security. A single platform reduces the overall architectural complexity because you don't need to have a software broker for some flows and multicast with WAN gateways for other flows, sometimes tied together with bridges and dedicated WAN optimization appliances between sites.

Application developers are more efficient since they can use a single API, talk a common topic and payload serialization structure without having to deal with the idiosyncrasies of multiple messaging products. This in turn provides significant architectural simplification and reduced time to market and risk for new applications. And as will be explained in this section, this can be achieved while addressing the requirements of each flow in a best-of-breed manner without sacrifices.

Price Flows – Non-Persistent Messaging

The distribution of pricing information demands ultra-low, predictable latency that remains flat even during spikes in volume.

Solace satisfies this requirement by implementing message processing in purpose-built cards built with high-speed network processors and FPGAs. Client applications connect to the message router using Solace APIs via TCP and subscribe to the topics they are interested in. Thanks to the use of discrete TCP connections they and receive only those messages since message filtering is performed in the message router, not with the client's CPU as in the case of multicast.

Delivery is achieved with microsecond latency between publisher and subscriber (API to API) and clients are isolated from spikes in traffic that they are not interested in. Clients are also not impacted by other slow or misbehaving clients because of the "shock absorber" function and per-client queuing performed by the message router.

Messages destined to slow consumers queue up in the message router and are delivered as the consumer is able to receive them, but publishers never need to resend messages, and other consumers do not receive retransmits destined for other consumers. This leads to a very stable real-time environment that is not subject to systemic collapse like multicast systems.

Need even higher message rates or more bandwidth? Solace customers can vertically scale their message processing capacity by changing a 2x10GE NAB for a 6x10GE NAB in their message router to access up to 80 Gbps of bandwidth (40Gbps in each direction) with low consistent latency, and all in the same datacenter footprint – without the complexity of horizontal scaling or needing to manage a multicast environment.

For delivery to desktop applications, the Solace message router provides authentication and authorization capabilities to allow you to pass all security audits. Use of TCP to desktops is by far preferred over multicast because desktop environments and networks tend to be much less controlled and thus require TCP to maintain a stable environment. Solace message routers implement functions such as intelligent message rate limiting so desktop applications are not overwhelmed yet always receive the most recent prices regardless of how slow they are and how fast the publisher is.

Solace also provides a last value cascading cache product that is accessed by Solace APIs to allow applications to retrieve the last value of a price as they subscribe – so they don't need to wait for the next price update. SolCache supports both top of book caching, typical in FX, but also depth of book where delta updates need to be merged in the cache.

Solace message routers implement various functions for efficient price distribution over the WAN directly in the message router – without the need for external WAN gateway processes or dedicated WAN optimization appliances. These functions include:

- **Streaming compression** is integrated in the hardware of the NAB to allow much more efficient bandwidth utilization and transfer far more data over the same network bandwidth. Compression ratios of 80% are very typical.
- **TCP optimizations**, such as those found in dedicated WAN optimization appliances, are integrated into the Solace NAB TCP stack. This provides full control over parameters such as TCP slow start and allows more aggressive recovery from packet loss. This type of tuning is not possible in standard server-based TCP stacks

- **Multiple TCP connections** are used between message routers to have multiple delivery streams, thereby increasing total throughput on long fat networks, yet while preserving order among related messages.
- **Dynamic routing protocols** ensure that the only messages sent to a remote site actually have subscribers to receive them. This is achieved by implementing dynamic, distributed routing protocols as is done by internet routers, only in this case for topic subscriptions
- **A dedicated TCP connection for control traffic** to ensure very fast network convergence when faults occur that is not impacted by the transfer of user data, resulting in a very stable system.

Order and Trade Flows – Persistent Messaging

Trades, and sometimes orders, require use of persistent messaging to ensure that these events are never lost. The key persistent messaging requirements for order/trade flow are high message rate, low consistent latency, slow consumer handling, high availability with fast, consistent failover and WAN distribution.

Message persistence is handled by the Assured Delivery Blade (ADB) within the Solace message router – which enables extremely high rate, low consistent latency and unparalleled slow consumer handling. As with the NAB, there are multiple versions of the ADB which allows upgrade options to provide vertical scalability.

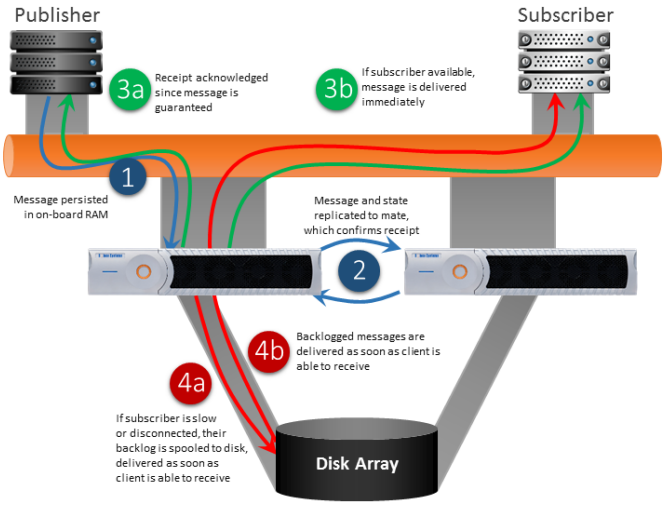
The chart on the right provides a view of the message rate at various message sizes for a popular software broker used in capital markets, for ADB2 and the higher end ADB3 blade. This rate is for a one-in, one-out (ie. No fanout) message exchange pattern.

This shows Solace persistent message rates are orders of magnitude higher than software message brokers. These measurements are with full “failsafe” storage of messages so they can never be lost. With Solace technology, there is no need to trade off speed for reliable storage of messages.

Low latency delivery of messages is achieved with Solace’s “cut through” persistence to allow for ultra-low latency delivery from publisher to consumer as shown in the following chart. Again, average latency is very flat as message rates increase and latency increases only slight in the 99.9th percentile as message rates increase.

The Solace message router allows consuming applications to select whether they would like messages received as durable or as non-durable at subscription time and without any impact on the publisher. Within durable subscriptions, consumers can select either cut through or store & forward delivery if they need more sophisticated delivery semantics.

Slow consumer handling is critical for persistent messaging systems as it is very typical to have both slow and fast consumers in a trading application (fast consumers in front office, consumers who cannot keep up with bursts in post trade) and because the message patterns are very bursty, the messaging system must ensure that neither it nor fast consumers are impacted by slow consumers – such as data warehouses, for example, which are often sized for average message rate not for peak rates.



High Availability and failure recovery times are also critical in trading platforms. Solace message routers use a patented high availability mechanism where message routers deployed in pairs provide an active/hot backup architecture, as shown in the following diagram, without needing any additional external software. All messages and message delivery state are synchronously stored (to ensure no loss) on the ADB of both message routers.

Upon failure of the active message router, the inactive message router automatically becomes active and all client applications automatically reconnect to the newly active message router. Because all messages and message delivery

state is already in RAM on the newly-active message router, service to client applications is restored consistently in a few seconds independent of the number or size of message payload stored for slow or offline consumers. In contrast, software alternatives must reload their messages and message delivery state from disk – a process whose performance varies based on the amount of data stored, but ranges from several 10's of seconds to 10s of minutes depending on the amount of data stored and the performance of the storage system. For a more detailed video on failover and recovery of persistent messaging, visit http://www.youtube.com/solacesystems.

Disaster recovery is also often a requirement where order/trade events stored in the messaging system must also be stored at a remote DR site. Solace provides integrated replication to a DR site without the additional cost, complexity and low performance of storage replication solutions. Refer to http://solacesystems.com/resources/disaster-recovery for a video on Solace's DR capabilities.

For persistent messaging over the WAN, the same techniques as mentioned for pricing (compression, TCP optimizations, and multiple TCP connections) also apply for persistent messaging.

Unified Application Connectivity Fabric

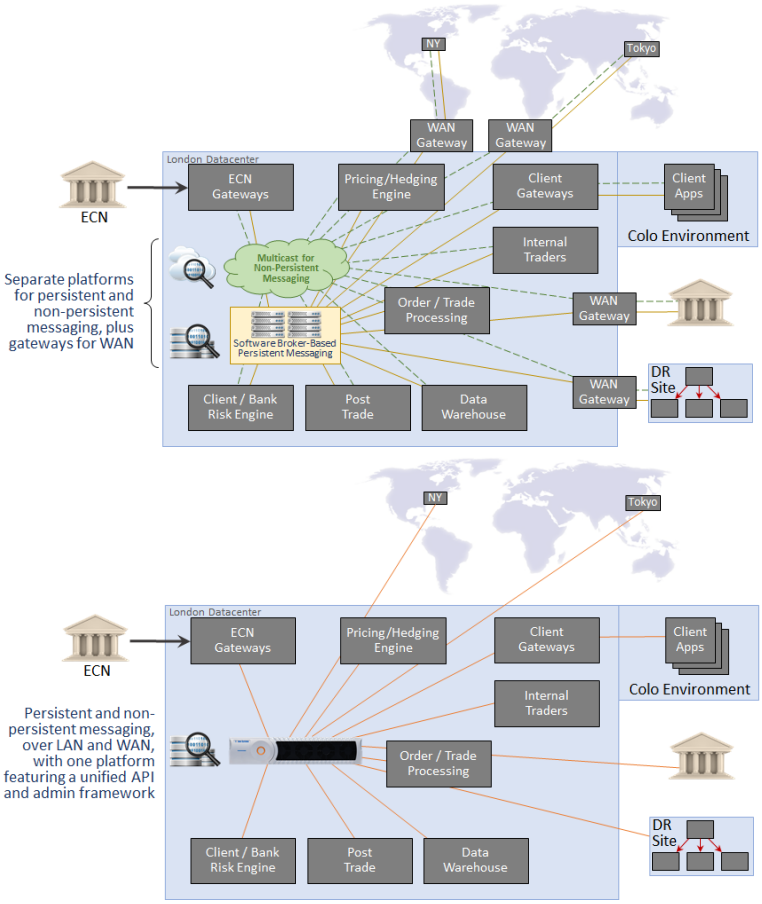
This pair of diagrams illustrates the difference between an FX trading platform built on software-based messaging and WAN optimization and one built on Solace.

Solace offers a single data movement infrastructure for architects and application developers to understand: one API, one topic structure, same mechanisms for all subscriptions, built-in application-level high availability, application striping for horizontal scalability, etc.

Solace also gives your operations team a single system to manage – one framework and interface for provisioning, monitoring, troubleshooting, capacity planning, and maintaining security.

Solace's WAN capabilities eliminate not just WAN optimization appliances, but also gateways between local and overseas destinations.

Solace offers the best performance without having to sacrifice loose coupling, robustness, manageability, system control or single-point-of-contact monitoring.



Management

Solace technology provides rich, detailed, real-time management – like the central nervous system connecting all components of your distributed trading platform. It is therefore often relied upon by clients as the one-stop-shop for monitoring the health of applications – something peer-to-peer systems cannot do and something software brokers struggle at because of the lack of separation of control plane and data plane, which is inherent in the Solace architecture.

Detailed performance and status information is available such as which clients are connected, what they have subscribed to, input/output message rates, current and high water queue depths, packet loss to/from each client and much more is

available in real-time for trouble shooting and capacity monitoring/planning. This information is available via Command Line Interface (CLI), via the Solace GUI (SolAdmin) as well as via a programmatic interface.

Events are asynchronously generated by the message router as clients connect/disconnect and as queues reach thresholds so various users are made aware of changing conditions. These events are emitted via SYSLOG and as messaging events on special topics for easy integration into your existing monitoring solutions.

All configuration information applied to the active message router is automatically replicated to the HA mate message router as well as to the DR mate message router in a remote datacenter to ensure simple and error-free configuration synchronization.

Security

Solace message routers provide centralized authentication, authorization and encryption capabilities that are separate for messaging applications and messaging administrators.

Access to the messaging services of the message router can be guarded using username/password authentication which is validated by LDAP servers (including Microsoft Active Directory) or by the Solace internal database. Access can also be restricted based on IP subnets to ensure that development applications do not inadvertently connect to a production message router even if the same username/passwords are being used.

Entitlements to publish to certain topics or subscribe to certain topics can be restricted on a per-user basis using the Access Control Lists provisioned on the message router. Violating the ACL rules causes the request to be denied and an event to be sent via SYSLOG and messaging.

Solace message routers support SSL encryption in hardware for extremely high speed datapath integrity protection.

Conclusion

Solace satisfies all of the data distribution requirements associated with powering world-class FX trading platforms, including the fast, reliable distribution of pricing data and resulting orders and trades. Solace does so with a unique hardware-based message router that is easier to deploy and operate, and offers superior performance, reliability and manageability. To learn more, visit http://solace.com.