

Improving Information Flow for Equities OMS & Trading Platforms

Within investment banks, equities order management systems (OMS) and front office trading platforms can vary in their requirements depending on the level of service the platform is providing. Some handle sophisticated high touch trades that require considerable human trader interaction, others take care of low touch trades powered by smart order routers, VWAP target or other algorithms, and direct market access and sponsored access trades involve no “touch” at all. In all cases, however, OMS and trading platforms need to handle large bursts of orders and resulting trades while maintaining ultra-low and predictable latency even when facing unpredictable and volatile demand from external client applications.

Regardless of the type of order flow, a post trade function is required to perform regulatory reporting, street and exchange side settlement, trade data warehousing and more. Many people now refer to this as the middle office since it sits between the extremely low latency front office and multi-day processing of the back office. The middle office “need for speed” is driven mostly by new regulatory reporting deadlines and the need for near real-time risk monitoring. The interface between the front office and middle office has different requirements than the front office, and is a critical component of the architecture.

This paper provides an overview of how Solace technology addresses the needs of OMS, front and middle office systems with a single platform while focusing on the main functions of each area, the information flows between them and the requirements of those flows.

Introduction

Solace has advanced the state of the art of messaging middleware by introducing message routers that can distribute massive amounts of information with different qualities of service and latency characteristics over a variety of network types. As a result, Solace can meet the unique needs of financial institutions that offer trading services by handling low latency communications between front office systems and guaranteed delivery from front office to middle office components.

For both front and middle office, Solace technology offers the following advantages:

- **Ultra-low, predictable latency at high message rates.** Solace technology offers very high performance non-persistent messaging rates (17 million messages per second) with very low, deterministic latencies (low 20 microseconds) even at the 99.9th percentile which is critical in the front office. Solace also offers very low and consistent latency when persistent messaging is required to handle trade. Clients can start with lower performance if needed and then vertically scale their system by upgrading to new cards to increase performance without increasing footprint and complexity.
- **High volume persistent messaging with fanout.** Especially for middle office applications, very high volume persistent messaging is required with fanout frequently around 5x the input rate. Events from the front office must be accepted at very high rate with very low latency and persisted so they are never lost, then often sharded (for parallel processing) and fanned out to a variety of middle office applications at a rate they can handle.
- **Slow consumer handling.** Trading platforms frequently see situations where one event is being received by several consuming applications – some of which can keep up with the bursts, others that cannot, such as trade data warehouses. Solace's unique slow consumer handling ensures that slow or offline consumers do not impact the performance of other applications of the message bus – a critical property to ensure stability of the overall system.
- **Integrated high availability with fast failover.** Solace message router have integrated HA which does not require any 3rd party software, thereby reducing complexity and increasing robustness. Solace's performance is constant even with HA enabled and switchovers occur in seconds regardless of how many messages or how much message payload is queued, thereby increasing availability and predictability.
- **Centralized real-time management.** Since the Solace message router connects all applications, it provides deep insight into all aspects of data flow and is often used as the central monitoring point of system sanity. From this single point, operators can determine what applications are connected, what they are subscribing to, what their queue depths are, what their input/output message rates are, who are the fastest publishers and consumers and much more. All this information is available in real-time without impacting application performance.
- **Intelligent congestion controls.** Since Solace message routers are in the real-time path of message exchange, they know the congestion state of all consumers and can thus influence real-time routing – for example to choose an uncongested exchange gateway toward a liquidity pool for a New Order to ensure lowest possible latency to reduce slippage. This is just one of many features purpose built into Solace message routers for front office applications.

Basic Trading Platform Architecture

This figure depicts the typical architecture of an equity smart order router (SOR) platform. It shows the front office system components of client gateways, SOR engines and exchange gateways involved in actual order execution. There are also typically trade management and monitoring functions passively observing events with the occasional need to support human intervention.

The top Solace message router provides high rate, low latency delivery of orders and trades between the various front office components. There is typically little message fanout in this case since most order/trade events are sent to a specific destination (specific SOR engine or exchange gateway instance) and often to a real-time trade management system, but not too many more destinations. Many front office systems also emit telemetry information regarding their processing to allow for white box application health status and performance monitoring/logging.

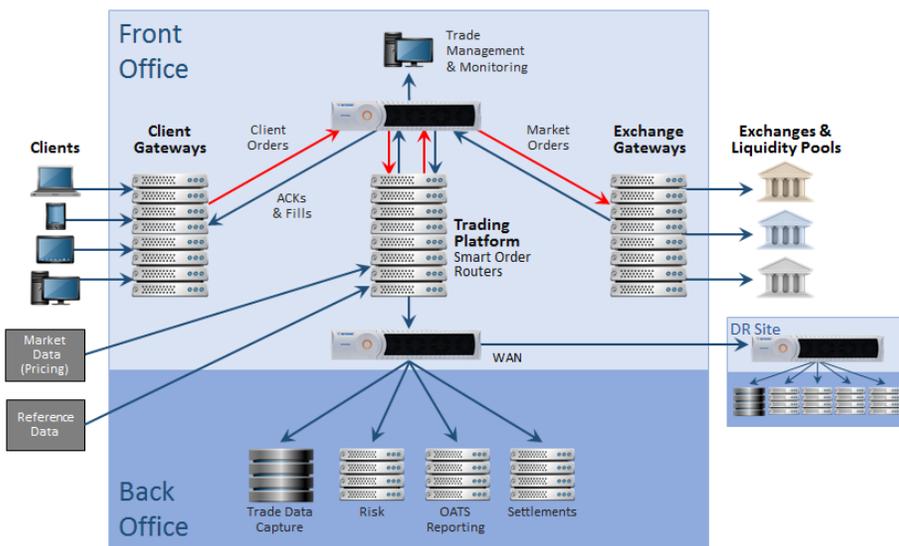
When crossing engines (dark pools) are supported, they can take on many different architectures such as having their own separate deployment like the one above (but without the Exchange Gateway layer) for general access by many internal systems, or crossing engines may be co-located with or embedded within the SOR function for lower latency access.

Front office systems are very often in co-location sites rather than inside the bank's datacentre for latency optimization reasons.

The platform also requires external information such as real-time market data from direct feeds and reference data on the securities it is trading.

Below the front office are the middle office systems such as Trade Data Capture, Risk Management, Regulatory reporting (e.g. OATS) and Settlements. Some of these could be physically in the co-location site but others are more often in the bank's datacentre. The interface between the front and middle office is provided by the bottom Solace message router in the diagram. For this middle office interface, only persistent messaging is used and various aspects of persistent messaging are critical:

- Order and trade events must be accepted from the front office systems at extremely high burst rates and persisted so they cannot be lost. The persistence must be such that under no conditions can it become slow to accept new trade events from the front office because if the persistence function slows it will impact the performance of front office trade processing or increase the risk of losing trade information.
- In this regard slow consumer handling is critical. It is typically the case that some middle office applications will keep up in at least near real-time with trade events as they occur (e.g. risk). However, other systems have no requirement to keep up with the bursts (e.g. trade data capture) and sizing these non-real-time systems to the scale required to handle the burst (vs. average) load would be very expensive. Therefore, even under normal conditions without faults there will be "slow consumers" on the message bus and these must not impact the performance of either the front office publishers or other middle office consumers. A real-time-to-non-real-time "shock absorber" function is required between front and middle office to handle this rate adaptation.



- It must provide fanout to multiple downstream systems, so message replication typically with a fanout of between 5 and 10 is required. The fanout must also be able to “stripe” the events to consuming applications while maintaining temporal order of events within the same order/trade to allow for horizontal scaling of middle office applications.
- It must be able to unspool stored messages very quickly to a consuming application that has built up a backlog (for example because it was down) and do so while continuing to persist newly arriving messages from the front office. This is necessary either to bring a crashed system back online and get it back into real-time, or to allow stored events to be delivered and processed by the middle office systems before a deadline (e.g. OATS reporting deadline) inside a smaller processing window due to an application or infrastructure outage.
- High Availability and fast switchovers from an active persistent messaging system to its backup without ever losing any messages is a clear necessity to minimize downtime in the event of a failure. This switchover time must not be dependent on the amount of data stored in the messaging system as the combination of slow consumers and variable failover times have a disastrous effect on platform uptime.

Applying Solace Technology to Equity Trading Platforms

Solace in the Middle Office

We start this discussion with the middle rather than the front office because the middle office interface is a critical component of the front office architecture but has different requirements and receives much less attention in technical discussions.

As described in the previous section, robust, high performance persistent “shock absorber” and fanout capabilities are the main requirements of the middle office interface – not ultra-low latency. This is why peer-to-peer messaging in particular is not well suited for capturing front office events and fanning them out to middle office applications.

The first requirement is the ability to support extremely high persistent message rates.

Slow Consumer Handling

Slow consumer handling is the next critical persistent messaging capability required in order to ensure the system is predictable since it is very typical to have both slow and fast consumers in the middle office.

High Availability and low failure recovery times are also critical requirements. 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 hot backup message router automatically becomes active and all client applications automatically reconnect to this newly active message router. Because all messages and message delivery state are already in RAM on the newly-active message router, service to client applications is restored consistently in a few seconds independent of the number of messages 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 our YouTube channel at <http://www.youtube.com/solacesystems>.

Disaster recovery is also often a requirement in the middle office 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. You can watch a video about Solace’s DR capabilities at <http://solacesystems.com/resources/disaster-recovery-video>.

Some sell side firms that have pre-existing use of a different messaging technology, such as peer-to-peer, in the front office, have found that these technologies lack the performance, stability and manageability for persistent messaging required for a

middle office interface. The requirements to never lose a message and to perform fanout to a number of consuming applications, many of which can't handle the bursts, results in slow consumer situations that are death for peer-to-peer architectures. As a result, they have decided on a dual technology approach – using Solace at the middle office interface point in the architecture due to its best of breed persistent messaging capabilities, while keeping peer-to-peer in the front office if this technology is working well for them. To achieve this, there are two strategies that can be used:

1. Create a bridge consumer off the front office bus that uses the peer-to-peer API to subscribe to messages destined for the middle office and then uses the Solace API to publish them to the Solace message router for fanout to middle office applications. This bridge application does little processing so it can deal with high bursts of traffic from the front office. The Solace message router then does the burst absorption by persistently storing all messages, then fanning and metering them out to middle office applications at a rate they can consume.
2. Modify front office applications to use the Solace API rather than the peer-to-peer API to publish messages to the middle office. This requires modifications to applications but avoids a bridge consumer on the peer-to-peer bus.

Which option is best is dependent on application architecture and sometimes organizational responsibilities.

Solace in the Front Office

In the front office, ultra-low latency is a critical requirement of the messaging system. However, not only low latency, but predictable low latency especially at peak message rates. Buy side clients count most on low latency during periods of volatility which in turn is when volumes are highest. It is in providing low, very consistent latency even at peak rates where hardware solutions shine. Solace's hardware based solutions use FPGAs and network processors and as such they have no interrupts, context switching, user-to-system space data copies etc. to create latency variability.

In some trading platforms, non-persistent messaging is used for new orders from clients whereas in all platforms persistent messaging is used for order acknowledgements and trade messages. For non-persistent messaging, the chart on the right shows the publisher to subscriber (API to API) latency for typical order messages using the 2x10GE and 6x10GE Network Acceleration Blades (NAB) at millions of messages per second – more than enough to handle new orders in an equity trading platform. Note the low, consistent latency even for the 99.9th percentile. Solace APIs also integrate with kernel bypass technology for the lowest possible latency.

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 and 40Gbps out of the message router) with low consistent latency, and all in the same datacenter footprint – without the complexity of horizontal scaling or needing to manage a multicast environment.

Trade and some order related messages require persistence in the front office so they are never lost in transit between application components yet they still require ultra-low latency delivery. For this, Solace message routers support a feature called “cut through persistence” which provides the full assurance of failsafe storage, but also provides ultra-low latency delivery. The chart to the right shows the low and consistent API to API latency for 600 byte messages (representative of trades).

The Solace architecture also provides high availability and fast, consistent failover times as explained in the previous section.

Architectural Advantages

A hardware broker-based architecture has several advantages over peer-to-peer messaging architectures:

- **Predictable Latency** - In terms of ensuring predictable system latency even under load, Solace's broker-based architecture ensures that client applications only receive the messages they have subscribed to – rather than needing messaging software to filter out unwanted messages on the receiving server. This means that applications and their servers are isolated from bursts of messages they don't care about. Solace's hardware datapath does not use interrupts, context switching or system/user space memory copies and is therefore not prone to the latency variability introduced by these software concepts. Since publishers are not impacted by slow consumers, they do not

waste processing time or I/O dealing with misbehaving consumers – a process that creates latency variability in peer-to-peer systems.

- **Increased Robustness** – Solace’s architecture ensures decoupling between publishers and subscribers which is critical for ensuring system reliability and avoiding complete systemic collapse to which peer-to-peer architectures are prone – specifically because there is no function managing traffic flows between applications. With Solace, hardware based filtering and queuing ensures isolation between applications in the case where some application (or its hardware) is misbehaving or slow. 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.
- **Intelligent Traffic Management** – Because the Solace message router is in the middle between publishers and consumers, it is able to support features purpose built for trading applications that aid in the real-time routing of orders during periods of congestion. As one example, in a SOR platform, New Order messages sent to a particular trading venue can use any gateway to that venue but once that path/gateway is selected, all subsequent events for that order (modify, cancel, replace) must follow the same path to avoid messages arriving out of order at the venue and causing incorrect results. So when sending the New Order message from the SOR, the Solace message router has features that help to avoid busy gateways to a given venue (which would introduce higher latency) and instead chose a less busy gateway to this same venue. However, other messages (modify, cancel, replace) would stick with the originally-selected gateway even if it happens to be busy to maintain message order.
- **Improved Management Visibility and Control** - Because the Solace message router is in the middle between publishers and consumers, it is able to provide information on which applications are connected, what they are subscribing to (to ensure all applications are up), and can also provide ongoing, one stop shop health information of the entire system. The message router provides real-time information on message rates to/from each application and amount of queuing to an application – all indicators of a well-functioning application or one that is in trouble. It can identify specifically applications that are overloaded or not performing properly to allow for proactive treatment. With multicast systems it is not possible to have this information – especially when applications are not behaving correctly – because there is no application-layer traffic visibility.
- **System Simplicity** – A Solace-based architecture is very simple to administer, monitor and control. It doesn’t need to be completely uniform in terms of switches, links and servers, and you don’t need to be an expert in debugging reliable multicast protocols and engineering multicast networks to manage it. The use of TCP and a message-level broker ensure a simpler environment to deal with, simplicity that leads to lower risk and better reliability.

In the past, the lack of performance and the unpredictable real-time behavior of software brokers made a broker based architecture difficult if not impossible to use in production, but thanks to the use of a hardware based data path, you can have all the benefits of a broker-based architecture (decoupling, robustness, single point of management & monitoring) without sacrificing performance or latency.

Solace also provides a last value cache product that is accessed by Solace APIs. SolCache can cache the very last value of a message for a particular topic, as is typical in pricing applications, or it can cache the last several messages related, for example, to an order so you can see events such as partial fills or fills of child orders for a parent order.

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. Refer to the Solace YouTube channel at <http://www.youtube.com/solacesystems> for demonstrations of SolAdmin.

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 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.

Conclusion

Solace message routers meet the demands of both front and middle office in equity trading platforms, and can also be used as a middle office interface even if clients are happy with their front office messaging technology. Solace's unique message router is easier to deploy and operate than messaging software, and offers superior performance, reliability, predictability and manageability. To learn more, visit <http://solace.com>