Building blocks of an exchange
Financial markets have always been influenced by changes in technology, but the current rate of change is unprecedented. These changes create new opportunities for investors and intermediaries, and present new challenges for regulators.

Technology is a great market enabler, especially for emerging markets, as it allows them to leapfrog the legacy structures found in developed markets. However, technological innovation benefits investors everywhere in terms of greater access and cost savings. Improved access to markets further democratises the investment process.

For hundreds of years, physical paper documents and human beings dominated securities operations. Technology has done away with much of the physical paper and now dominates markets everywhere. However, technology and communications have transformed market structure driven by a tsunami of data which needs to be disseminated and analysed.

Sophisticated IT allows regulators and market operators to make sense of this high-speed and high-volume marketplace, and for the venues themselves to gain insight into how data is being consumed.

The volume of data is set to increase further as the expansion of internet connectivity to an ever greater number of devices - the ‘Internet of Things’ - creates a requirement for feedback control loops, increasing messaging output considerably.

While operational data provides increased control, data is also the raw fuel for indices. The business of indices for exchanges is a core strength and a key driver of growth. Technology allows index providers to bring indices on board quickly to meet market demand or to respond to competitive threats and the challenges posed by increased regulatory oversight of indices and benchmarks.

To operate efficiently in the fast moving interconnected marketplace of today, market participants are asking for technological solutions which are flexible, innovative and cost effective allowing them to evolve within the global markets ecosystem as regulations and technology develop.

“The business of indices for exchanges is a core strength and a key driver of growth.”

Herbie Skeete, managing director, Mondo Visione
Introduction

Technical building blocks of an exchange

Exchanges are venues that bring together buyers and sellers to effect a trade. Central to an exchange is the trading system on which trades are executed.

All exchanges have rules to ensure the integrity, efficiency and fairness of the marketplace and have a department responsible for managing all aspects of regulation relating to the marketplace. This will normally have sophisticated surveillance tools to maintain an orderly market through real-time monitoring of trading patterns that could indicate breaches of the exchange or regulator trading rules.

In today’s world, trading in many instruments is largely automated, providing greater scale of action. Reaction speed is determined by latency, not wit. Error and manipulation can also be scaled up. Systems are tested to destruction and are housed in specialised environments engineered for low latency and exacting reliability standards.

The trade cycle is made up of the following activities: pre-trade decision, order routing, order execution, matching, clearing and settlement. As crucial as the pre-trade and trading itself are the post-trade processes of clearing and settlement. After a buyer and seller effect a trade in a financial instrument, only when instructions from the counterparties are matched up, and checks made that the securities and cash are held by buyer and seller is the settlement process is completed.

For the pre-trade decision to take place one needs to know the prices prevailing in the market for the financial instruments being traded. Market data, the bid and offer quotes, last trade and volume information distributed by exchanges in real-time or delayed, either directly or through specialist information providers. Market data is one of the key sources of revenue for an exchange and is also a means of attracting liquidity to an exchange.

Specialised network providers bring together trading partners from around the world. Such networks have ended the tyranny of distance so that exchanges can now be local, regional or truly global.

In this market overview primer we, through our partners, will illustrate the processes outlined above.

Trading engine

Central to a trading venue is the trading platform, typically a central limit order book on which counterparties’ orders are matched. Market fragmentation and increased market complexity mean that for trading platforms to be fit for purpose they need to be scalable, with high reliability and capacity. Above all trading systems need to cope with the volumes and transaction speeds of today’s markets, and properly implemented and designed if we are to avoid more ‘flash crashes’ and market glitches.

Testing

Markets still worry about technology glitches after the May 2010 ‘flash crash’. Regulators and market operators in reaction to concerns about the strength of critical market infrastructure are demanding that measures be put in place to reduce the occurrence of systems issues and improve resilience when systems problems do occur. Resilience is provided through system architecture but also through thorough testing of every element in the system through every path. Thorough testing demands not only excellent technical knowledge but also domain knowledge of the markets.

Indices

Indices measure the performance of a market or sector, providing the underlying basis for many exchange-traded products, allowing the market that provides index data to profit from the data supplied to an asset manager offering an exchange-traded fund. Exchanges either design or manage indices themselves or else they collaborate with an index group to manage the indices and to maximise the reach of the indices to global investors. The integrity of benchmarks and indices is vital to the global financial system.

Post-trade

Post-trade services are crucial for the functioning of an efficient financial system effecting the transfer of ownership and cash after a trade has been made. Technology has effected fundamental changes in the infrastructure of the financial services industry driven by regulation and the use of emergent technologies such as the application of blockchain technology.

Surveillance

Trading venues have a responsibility to maintain an orderly market through real-time monitoring
of orders, trades and quotes to identify market manipulation, insider trading, layering and other breaches of the exchange trading rules or regulations. Technological advances have significantly affected the global securities market accelerating the movement of capital across international borders markets more integrated. These technological advances have made the job of maintaining fair and orderly markets more difficult, but at the same time have given rise to innovative solutions to policing markets.

**Market data**

The market-data industry is going through unprecedented change. Irrespective of whether you are a classic information vendor, a trading venue, a broker or a niche information provider, this change will affect your business.

New technologies, new business models and regulation are some of the disruptive forces at work. Market data consumption patterns are also changing. This impacts trading vendors and information vendors alike.

There is an inexorable trend towards voracious consumption of data by machines rather than being viewed on screens. These machines require richer data from a greater number of sources and across more asset classes.

The key market data dynamic is the bundling of analytics and data is creating a fundamentally complicated ecosystem which will allow numerous small firms to flourish. This will mean that in the medium term the large dominant players will need to right size. How they do this will be one of the more interesting business stories in the coming years.

**Apps**

In today’s world of interconnected market-places, data dominates. Digital data is part of every aspect of our markets. Making sense of this new reality is challenging, we need flexible yet cost-effective solutions for multi-asset trading, portfolio analysis, decision support, compliance, market data and all the other components interconnecting global financial markets.

**Business intelligence**

To have information, but to be unable to use it, is a frustration that many businesses experience. Freeing that information, and making it usable alongside other measures, can present real opportunities. Exchange operators and post-trade service providers face rapidly evolving business dynamics and competition. Seizing opportunities is therefore immensely important and to do so market operators need to see a fuller picture using the data they already hold.

**Hardware**

For the past 50 years computing power has doubled every two years. This trend, which is dubbed Moore’s law, is transformative because it means computing power is increasing exponentially. Markets have over the last decade experienced a revolution driven by the Moore’s Law rate of change provided by technological innovation and through regulation.

**Data centres**

In today’s environment where financial institutions have to become more agile as well as to keep tight control on costs, trading venues often choose to outsource their data centres to specialist providers who can provide state-of-the-art but cost-effective technology. Today’s high-tech data centres bring together trading venues and market participants in a low-latency environment, supporting the evolving needs of the global financial services industry.

**Networking**

Financial markets are constantly evolving, and market participants are always looking at ways to efficiently access new counterparties, geographies, liquidity venues and trading services. Market participants need to quickly connect and engage within the financial ecosystem. Trading venues whether local, regional or global need to connect to buyers and sellers. As markets become ever more interconnected, connectivity between a trading venue and its clients is crucial. MV
The matching engine
by Feroz Cader, vice president, Head of Trading Systems, MillenniumIT, Technology Services LSEG

The matching engine is the core of any trading venue. In its simplest form, it facilitates the electronic matching of orders from buyers and sellers, in a fair and orderly manner, within a regulated framework. This simple function however, sits at the centre of an increasingly complex and competitive global financial system. It brings about many interesting challenges; whether you are setting up a brand new exchange in an extremely competitive developed market, upgrading the legacy systems of an existing exchange in a developed market, or introducing electronic trading for the first time in a frontier market. Building the matching engine requires skill as much as it requires experience and capacity to deliver. It is a highly specialised niche requiring significant engineering expertise in software design, operability, latency management and connectivity, combined with a deep understanding of trading across multiple asset classes and risk management capabilities.

Availability and quality at source
The matching engine is mission-critical software that global financial markets depend on to provide participants with the confidence to execute their strategies. It must remain glitch-free and provide as close to 100% uptime as possible in a competitive environment while at the same time being flexible enough to respond to regulatory change in near real-time. Managing that change is much akin to replacing the wheels of a moving truck while keeping the truck on the road.

The ability to deliver such quality and availability starts at software design. Quality-at-source and fault-tolerant architecture are critically important to keep the cost-of-ownership low and to minimise opportunity costs from a time-to-market perspective.

Operability
Even a defect-free system is prone to operator error. A typical reactive solution is automation. However, there is no such thing as 100% automation and the operations cycle is not limited to the start and end of trading. It begins at the point of release deployment into production. Operability should be a key architectural building block and not an afterthought. A single permissions-controlled dashboard that provides a view and control into the system for reference data, client connectivity, matching and market-data across multiple markets and asset classes should be a standard operability feature along with automated release deployment and rollback capability.

Multi asset and multi-micro market
A matching engine should be capable of trading multiple asset classes in different market structures on the same platform. This design feature allows for harmonisation across platforms and other systems from pre- to post-trade. Even if a trading venue operates multiple segregated instances for different asset classes, a matching engine’s capability should allow a common code path to be deployed across these systems. This can significantly reduce maintenance costs compared to maintaining separate systems for multiple asset classes.

This is equally true for frontier or emerging market exchanges where running separate matching engines across multiple asset classes has proven to be prohibitively expensive, especially when such markets are in the early stages of launching asset classes beyond cash equities, and as such, trade very low volumes.

“A matching engine should be capable of trading multiple asset classes on the same platform.”

Increasingly, new regulations are demanding greater transparency for asset classes that are traded over the counter (e.g. FX, IRS etc.). Having the capability to on-board new asset classes with different trading models (micro-structures) on the same system will bring cost savings to the organisations leveraging this opportunity.

Latency
The arms race to be the fastest may arguably be over. It
is no longer a significant differentiator given that most systems are capable of microsecond levels of latency. However, to a tier-1 market, consistency in latency is an important factor and in most cases still a challenge due to latency peaks or outliers. Reducing outliers will limit the uncertainty faced by trading algorithms and in turn minimise the risk of algorithms misbehaving. Improving latency can help trading venues do more with less in terms of the hardware footprint and risk management, even if single digit microsecond level latency is not important.

Connectivity
A key aim of emerging and frontier capital markets is increasing liquidity, but barriers to entry for global investors tend to be high. It is important during a market’s evolution, to ensure technology facilitates rather than prevents increased investment inflows. Industry standard connectivity both at order entry and market data level is critical to ensuring the matching engine allows low cost and easy access to the trading venue when hurdles to foreign ownership are relaxed.

The fight for liquidity is not limited to emerging and frontier markets. In mature markets, very few trading venues operate in a single market environment. Price formation is decentralised and as such, market participants need the ability to connect to multiple exchanges to provide best execution. They can do that by maintaining individual connections to each trading venue but this can prove costly given the rate of change in each market. Trading venues should instead look to provide a single standard interface and shield its participants from the need to maintain multiple order and market data interfaces. To do this, the gateway framework of a matching engine needs to be multi-venue capable while maintaining consistent microsecond-level latency without compromising functionality of any connecting venue. Further, the ability to normalise connectivity across venues will be a differentiator with developed as well as emerging markets given an increasing trend to create regional cross-border trading partnerships between venues.

Risk Management
A matching engine’s ability to provide risk management should be a key consideration in the current environment of risk regulation being imposed on trading venues. There are two types of risks that can be managed by matching engine software, namely price/size errors (also known as fat finger errors) and exposure-related risks.

A price- or size-based error could negatively impact a fair and orderly market. A trade with an erroneous price or size could cause market-wide panic and adversely affect trading algorithms and valuations. At a minimum, matching engines should be able to manage this risk through price bands and order size checks.

“A trade with an erroneous price or size could cause market-wide panic and adversely affect trading algorithms and valuations.”

Exposure risk is the risk of a default. Regulation and the complexity of unwinding a trade, post-execution, is shifting the need to perform risk management upfront, while latency improvements are making this possible. A matching engine should have the ability to be ‘participant risk aware’ by integrating with a risk management system to maintain real-time risk thresholds per participant without compromising on its matching performance.

Experience is key
Competition among exchanges for market share and new regulations are driving change at trading venues at exceptional speed. Deploying a matching engine is labour intensive. It is vital that a matching engine provider has both scale and a proven product, along with experience and process efficiency to support a trading venue’s evolving requirements. As more venues turn to specialist providers they can benefit from sharing the cost of development across multiple organisations at a fraction of the cost, while accessing innovative ideas from multiple sources and simplifying member connectivity by leveraging standard protocols. MV
The modern society expects quality and reliability from the latest exchange platforms. Thorough verification of exchange technology is not possible without adequate software testing tools. Understanding their evolution can help deliver rapidly evolving advanced exchange systems.

Test automation tools allow simulation of connected automated systems, running tasks in an unattended mode and generating load as in real-life market environment. The process of selecting a tool to test a trading or post-trade system usually starts with looking at generic testing solutions widely used in other industries. No one has been fired for using standard commercial solutions, but the problem with most of them is that they are targeted at desktop and web interfaces. However, integration with trading APIs usually requires developing connectivity and session libraries as well as a framework that would allow to uniformly define message contents and expected results. Essentially, one will need to build one testing tool on top of another.

Exchange testing tools evolution

There are a few specialised testing tools available that already support FIX, FAST, ITCH, binary exchange protocols and have expandable architectures. The user specifies what to send into the market and what the expected result from every gateway should be. This process is called active testing. If the tool fits the purpose and the QA team is capable enough, thousands of test cases to validate most of the exchange platform’s functionality under test can be prepared and maintained.

The execution of huge regression libraries produces massive amounts of data. Additional insights can be obtained from analysing these large data sets. The next evolutionary step in test tools is passive testing where one uses non-intrusive test tools that capture all messages generated in the system, normalise them, and store them for processing. Passive testing tools can help obtain valuable results about the system and the test library coverage. The Exactpro team has also found that passive test tools are invaluable in supporting client on-boarding and certification. Clients connect to the exchange platform and perform the necessary steps, while the tool captures these activities and produces pertinent client self-certification reports.

One can create a good functional testing tool and a scalable load-testing tool in order to generate hundreds of thousands of messages per second and stress one’s platform to its limits.

However, there is another area in between—intermittent defects that do not manifest themselves every time the tests are performed, potential black swans. Such defects are at the confluence of the functional and non-functional testing. To trigger them, one needs to repeat the tests many times and under load, and they can be easily missed in the enormously large load testing logs. These defects might not have severe technical consequences, though.

Passive testing tools can help obtain valuable results about the system and the test library coverage.”

Ideal passive testing tool

What characteristics should a passive tool have to capture defects at this point? It should be scalable, stable, have minimal impact on the system under test; it should collect all events from all components in the system; it should efficiently capture all of the data for analysis; it should notify the user when something goes wrong; it should support the ability of the user to configure scenarios that trigger such a notification; finally, when something indeed goes wrong, the user should be able to restore the transactions around the event.

An ideal passive testing tool for exchanges functions exactly like a production monitoring / market surveillance solution.
Is that it? Can the test tools of the next generation be just algo-trading and market-surveillance platforms? Not really. There are two distinct reasons for that. First, imagine trying to fit several algo engines and an additional monitoring solution into your budget. It will substantially increase the cost of testing. The other limitation in using ordinary real-life systems is that their intention is to do what is right. However, the purpose of testing is to find what is wrong. Testing tools of the next generation will be very much like production systems, but deformed and mutated to decrease the hardware footprint and to enable all possible permutations.

In a more distant future, there might be test tools based on machine-learning techniques. However, they can be enabled only if the test tools of the next generation are scalable and mutable. As singularity is not here yet, human intelligence is still required to create and operate test tools and large regression libraries. The best part of any testing solution is the human mind behind it. Exactpro builds software to test our clients’ software and has developed tools for trading, clearing and surveillance systems testing. We constantly work on evolving our test tools, so if one ever needs a testing tool of the next generation and/or a comprehensive quality assurance process, we can help.

"As singularity is not here yet, human intelligence is still required to create and operate test tools and large regression libraries. The best part of any testing solution is the human mind behind it."
International investors’ access to the domestic ‘A-share’ segment of China’s equity market was heavily restricted until the introduction of the Qualified Foreign Institutional Investor (QFII) programme in 2003. In the first few years after the launch of the programme, China’s regulators granted QFII quotas at a slow and steady pace.

However, in the past few years, China has accelerated the liberalisation of its domestic capital markets. The launch of the Renminbi Qualified Foreign Institutional Investor (RQFII) programme in 2011, the simplification of the R/QFII application process, the rapid growth in the approval of R/QFII licences, the extension of the R/QFII scheme to other asset classes and the resulting expansion of the aggregate R/QFII quota are all testaments to this policy shift.

As of March 2015, 414 QFII and RQFII licenses have been granted, a 25 per cent increase compared to last year, and the total quota approved increased by 46 per cent, reaching US$ 125.3 billion.

Source: FTSE Russell, CSRC, SAFE, data as at 31 March 2015
A further step in the liberalisation of access to China’s A-share market came with the introduction of the Shanghai-Hong Kong Stock Connect programme in November 2014. The programme permits international investors with Hong Kong brokerage accounts to access A-shares and mainland investors to access Hong Kong-listed shares, subject to separate quotas (initially US$ 47.9 billion for ‘northbound’ trading and US$ 39.9 billion for ‘southbound’ trading). The programme is open to both institutional and retail investors.

A-share benchmark solutions for international investors

Despite this trend towards the liberalisation of China’s domestic capital markets, A-shares are so far ineligible for inclusion in FTSE’s Global Equity Index Series (GEIS). This is because China A-shares market does not yet meet the requirements for market accessibility, free capital repatriation and clearing and settlement cycle that are governed by FTSE’s Country Classification Committee.

Nevertheless, many international investors have now received substantial quota allocations to enable them to access the A-share market and, increasingly, require index solutions to reflect their levels of access. As a result, FTSE has launched the FTSE Global China A Inclusion Index Series. This index series has two primary objectives: to provide a comprehensive range of benchmarking solutions for these market participants and to prepare for the potential inclusion of A-shares in FTSE’s standard indexes.

The index series also aims to serve as a benchmarking tool for international investors wishing to access China A-share market, to reflect the prevailing levels of market access for international investors and to be transparent and easy for users to understand.

The FTSE China A Inclusion Indexes provide market participants with following benchmark options:

- Global benchmarks with China A-shares included and weighted by the aggregate approved quota (including QFII/RQFII);
- Global benchmarks with China A-shares included and weighted by free float and foreign ownership-adjusted market value, assuming no quota restriction;
- Customised indexes based on an investor’s own QFII/RQFII allocation.

The index series follows the existing FTSE GEIS methodology, where China is reviewed as a single entity.

In the diagram we show how the FTSE Global China A Inclusion Index Series is constructed by adding the China A-share market available to international investors to the FTSE All-World and FTSE Global Small Cap indexes.

FTSE has a long-standing presence in China and in 2015 celebrates the fifteenth anniversary of its first Chinese equity indices, launched in 2000 under a partnership with Xinhua Finance Limited.

In 2010, FTSE acquired full ownership of these indices, which were renamed the FTSE China Index Series. FTSE’s China indices have become the benchmarks of choice for institutions worldwide and underlie a wide range of index-tracking mutual funds, ETFs and derivatives.

As the leading index provider for China, FTSE maintains offices in Hong Kong, Beijing and Shanghai, supporting clients in the region with a full suite of research, distribution, support, product issuer and asset owner services.

As at end-December 2014, ETFs with US$24.4bn in assets worldwide used FTSE China indices as their reference performance benchmarks, including ETFs on the flagship FTSE China 50 and FTSE China A50 indexes. MV
For market operators, the ability to derive real-time market intelligence from various forms of data is an essential ingredient for success. Internally, it helps them understand what is working and what is not, and it allows them to be proactive in making improvements. Externally, it helps them understand their customers, which leads to deeper relationships and better services.

Imagine being able to provide senior management with up-to-the-minute reports following a new product launch. Just think of the scope for improving satisfaction by sending alerts when customers’ order-to-trade ratios are nearing a point where they will breach thresholds.

In reality, market operators often struggle with the real-time analysis that underlies market intelligence. Today even smaller venues must cope with enormous volumes of historical and streaming data. Since data is recognised as a strategic asset, they are collecting and storing more than ever before. Reference data, market data, trading data, account data and counterparty data all reside in their core systems and various silos, but it is in several formats. It must be extracted, cleansed, normalised and made accessible to an array of business applications in a cost-efficient manner. It is important to capture the data in its most granular form to avoid loss of quality.

The goal is to take raw structured and unstructured data and turn it into information, knowledge and wisdom, and ultimately monetise it – and that is not an easy feat.

Executive need to base strategic decisions on a complete picture, but data alone does not provide that. Relevant pieces are scattered across the data in various shapes and forms, making them hard to find. When data is structured and ordered, it becomes documented, accessible information. But even then, information does not automatically lead to knowledge.

Events and information that are stored in disparate systems can be turned into knowledge by applying analytical models and exploratory visualisation to correlate them and find hidden patterns in the data. Visualisation helps market operators quickly identify anomalies and outliers in large datasets. When visualisation is overlaid with business knowledge and expertise, market operators can gather new insights that enable strategic decision making and accelerate business growth.

This can be done without sacrificing quality or incurring excessive costs. An open data management framework can be implemented at the foundation of the enterprise architecture. As such, it can support all core systems that enable the trade life cycle, from the order through to trade, clearing and settlement. Essentially, the data management framework serves as an ‘information bus’ that collects all the data from various operational systems and external sources of relevance, and transfers it to a centralised infrastructure where it can be managed. Since it is ‘open’, the market operator is free to use its preferred method of data storage and add applications as necessary.

Think of this as a paradigm shift from the
Traditional data warehouse concept where data is at rest: it is simply loaded, transformed and extracted, and decisions are made based on stale data. In the data management framework, however, data is in motion. Live structured and unstructured data from the markets, internal business systems, social media, mobile applications, the news and other sources are fed into the framework and combined with historical/static data. Then it can be sliced, diced and refined using complex event processing. It can be sequenced and replayed. Market operators can run real-time streaming queries and combine historical queries with streaming data. Calculations can be made on the fly for predictive analysis.

The framework provides a unified API for integration, ensures relational integrity and resilience, and stores high velocity data without delays. Importantly, the data can be tapped and collected without affecting other systems. When the data is visualised on dashboards, executives can make more effective decisions.

Connecting data and allowing users to see it in their own context can be incredibly valuable. If data is correctly articulated it can reveal patterns that were previously hidden and make otherwise unknown relationships visible.

Consider the following examples:

- Data from a trading engine, the billing system and an external market data feed could be aggregated and cross-referenced to analyse performance. Market operators could identify a shift in demand for specific products and determine the cause. They can identify which products generate the most revenue and attract the most liquidity. Armed with this market intelligence, they can make strategic decisions about product design and fees, and determine the optimal way to allocate resources.

- Historical data can be cross-referenced with order and trade data to determine the cause of increased latency at a specific time. With this insight, market operators can have an informed dialog with participants to help achieve an optimised experience and increase customer satisfaction.

To compete in today’s data driven, globally interconnected environment, information is key to gaining an advantage and staying agile and relevant. Moreover, global regulators are requiring various data sets to be made available upon request and often within tight timeframes. While accurate, accessible data is an undeniably crucial component, obtaining that data in a timely manner – especially across multiple formats and sources with infinitely increasing volumes – continues to be a challenge for many market operators. Startup marketplaces can avoid falling into this trap by putting a data management framework on the IT roadmap from the beginning, while incumbents can solve an existing problem by adding it to their enterprise architecture.

Market operators can start by implementing a data management solution and collecting all data in its most granular form. Before long, they will realise the hidden value in the data and begin applying data management techniques in various new ways to gain business insights they never envisioned before. Finally, they can leverage data to develop innovative offerings and expand customer value while potentially creating new revenue for the marketplace.
The financial crisis revealed that the industry’s approach to risk management was flawed. Since then, new regulations have been created to push more activity onto regulated marketplaces and central clearinghouses (CCPs), which are deemed safer than bilateral trading and clearing arrangements.

CCPs manage, mitigate and mutualise risk by serving as the counterparty on all trades (the buyer to all sellers and the seller to all buyers). The members put up capital in a mutualised guarantee fund. They also post initial margin when they take a position, and variation margin when the market moves against them. If a firm cannot meet a margin call, its positions are liquidated. When necessary, CCPs can tap into their shared capital and their own capital.

By clearing through a CCP, firms can benefit from multilateral netting, where several parties’ transactions are summed instead of being settled individually. Multilateral netting not only reduces credit, operational, settlement and liquidity risk, but it also reduces firms’ balance sheets. It is a cost-efficient mechanism for managing initial and variation margin because firms have one net position and one collateral pool to cover their margin calls. Central clearing also leads to a reduction of regulatory risk capital charges.

CCPs can drive operational efficiency by enabling members to optimise collateral management using real-time information, straight through processing and the auto allocation of funds on a client level. They can help members identify surplus collateral, free up valuable liquidity and maximise value from securities finance transactions.

Through portfolio margining (combining positions in different products to offset margin), CCPs can help members reduce their overall margin requirement. Depending on the degree to which collateral optimisation services are effective, there may be significant reductions in the amount of overall collateral required.

Central clearing enables participants to trade and clear with everyone without having to worry about direct counterparty risk. CCPs have a lower risk profile than traditional banks because they always have balanced positions, except in a default situation. They are in the business to manage risk, not to speculate. Moreover, they must comply with regulations that require certain levels of capital and risk management processes. Markets are safer as a result, which attracts more global participants, boosts liquidity and drives competition on exchanges.

Four key components are needed to build a CCP.

**CCPs have a lower risk profile than traditional banks because they always have balanced positions, except in a default situation.**
The first is proven and tested technology. CCP technology needs to be flexible because both startup and established organisations need room to grow and add new products as they move toward a multi-asset, multi-currency and real-time environment.

After a trade is executed, the trade title is transferred to the CCP. To this end, they need a system to handle novation (CCP matching) and netting. They need a real-time risk management tool so they can constantly evaluate their risk. In the event of default, CCPs must be able to close out a portfolio at the market prices assumed in their model.

To do portfolio-margining, CCPs must be able to run a series of rigorous calculations to identify which products are correlated, demonstrate the strength of the correlation, and determine whether the correlation is stable over time. They also need to monitor performance and back test to ensure that the correlation will hold up in current market conditions.

CCPs need a system to hold various collateral assets including bonds, equities and cash. It must be possible to segregate collateral in customers’ accounts so their assets are protected should a clearing member default. By using real-time prices and positions, CCPs minimise the margin requirements they impose on their members and customers.

Data management is critical. Customers want to trade and clear more products faster than ever before, which introduces a new level of complexity, especially in margining. CCPs need graphical user interfaces and gateways to aggregate data flowing in from various exchanges and liquidity pools. New, updated and historical data enables them to perform sophisticated risk calculations that determine margin levels, and influence the degree of leverage and the amount of capital to be committed to positions.

Additionally, CCPs need various systems to handle treasury functions, reconciliation and corporate actions.

Second is regulatory status that delivers capital efficiency and complies with global standards for CCPs while meeting local or regional demands. CCPs are systemically important financial institutions, so they must be certified to ensure they meet regulatory standards. In most legislations they also must be set up as separate legal entities and operate as standalone companies with their own board and capital.

Third is competence and know-how. Although central clearing is less risky, it is not risk free. It is critical that the CCP is operated correctly with appropriate processes, and it collects the right amount of margin and default fund contributions. Only products that can be properly risk managed and liquidated in a default scenario should be put into the CCP.

Fourth is the CCP capital structure. CCPs have a waterfall mechanism whereby various levels of safeguards protect against the default by one or more members. The waterfall should include both participant and CCP capital and be designed to create the right risk incentives to ensure a functional default management process.

Members post margin with the CCP, and the CCP has a default or guarantee fund. If the guarantee fund levels are higher, then the margin levels tend to be lower. CCPs model and analyse the size, nature, diversity and volatility of their members’ portfolios so they strike the right balance between margin and guarantee fund levels.

The risk management framework of a CCP comprises several layers, all of which are fundamental to building a strong organisation. A few best practices should be kept in mind.

When building a CCP, it is critical to develop prudent member admission criteria and ensure it is implemented. Margins should be calculated to cover for the market risk inherit by the CCP. CCPs should set strict guidelines and haircuts for the types of collateral they are willing to accept. Finally, they should stress test participants’ portfolios, strike a balance between margin and guarantee fund levels, as well as determine how the pieces of the waterfall fit together and when margin levels should change.

“Only products that can be properly risk managed and liquidated in a default scenario should be put into the CCP.”
Powering exchange integrity

By Robert Lang, vice president, head of Product Management, Risk and Surveillance Solutions, Nasdaq and Tony Sio, head of Sales and Account Management, SMARTS Market Surveillance, Nasdaq

Fair markets are free from manipulative and deceptive behaviour, with clear rules that are enforced fully and without bias. Orderly markets are those which are reliable, without unreasonable price fluctuations or errors. Without these characteristics both investors and potential issuers will stay away, starving the exchange from both directions. An effective surveillance program plays a key role in every exchange’s future.

Recent changes in technology, regulation and global economics have revolutionised the securities markets. Dynamic exchanges have adapted quickly and launched new products and services to meet new customer demands. But to remain successful, exchanges must also ensure that their surveillance program keeps pace with change and that they have the same abilities to view the market as the traders they are monitoring. The minimum requirements include a full consolidated audit trail and real-time, automated and cross-product/cross-market monitoring.

Cross-market, cross-product surveillance

Globalisation and fragmentation allow market participants to seek opportunities in multiple assets and geographies, and take advantage of differences in structure, pricing, latency, hours and order types – exchanges need to be able to detect these advantages. The effectiveness of a single market’s surveillance efforts is reduced if the system cannot incorporate trading data for the securities across all the venues on which those securities trade and their related instruments. For example:

- Layering (Spoofing): Analysts may be alerted to layering of the order book on one market, but without cross-market functionality, they may not be able to identify that the same account is trading on the opposite side of the market from the layered orders on another venue.

The markets are so inter-connected that surveillance cannot be done in a siloed manner. The economic relationship between the traded instrument and related instruments must be considered to identify gaming and manipulative behaviour that may occur between venues. A participant may try to manipulate the price of a particular stock in order to make a profit on an options contract on that stock. Alternatively the manipulator may be profiting through over-the-counter contracts or other linked products.

In addition to possibly missing market abuse, compliance staff may be wasting time and resources investigating scenarios based on single venue, single asset class data. Behaviour that appears unusual in the context of one venue may be normal when viewed against the backdrop of multiple markets or related instruments. Hence, the ability to view consolidated trading data from multiple sources in a single view has become essential, as has the ability to review alerts in the proper context, and compare them to what is happening in electronic communications, news and other areas.

Order-based analysis over execution

Historically, surveillance has focused on market integrity and market orderliness separately. Rule breaches as well as unlawful and dishonest conduct relate to market integrity; errors, system stability and market activity relate to market orderliness. The ability to reconstruct every aspect of the market at any point in time, and at the very lowest level of granularity, is the foundation for monitoring both areas. Trade and top-of-book data is not sufficient to accurately understand modern surveillance issues as many types of manipulation cannot be detected at the trade level. Regulators are devoting more attention to order-level manipulation. With order-level analysis you can get closer to determining true intent.

Real-time control

Exchanges cannot wait until the end of the day or the week to analyse trading activity.

While historical analysis is still extremely important in determining abuse, analysts need to be alerted instantly when a market event occurs. A failure to take action could unleash extreme volatility, spread contagion to other markets, cause huge losses and erode investor confidence. A prime example is
the Knight Capital case, in which the company lost about US$170,000 per second over the course of 45 minutes. Real-time capabilities are critical to early detection of unusual trading patterns that could be potentially catastrophic breaches of trading rules and practices.

**Technology**

Modern surveillance technology uses data mining and pattern recognition techniques to monitor the markets. First, it combs through large amounts of historical data and builds a profile of normal, day-to-day behaviour. Then it captures trading activity in every product listed on the exchange in real time, and compares it with the norm to detect certain types of events. Anomalies can be uncovered quickly, and exchanges can immediately make a decision to launch a preliminary investigation or even halt the market.

**Data integration:** Storing and handling data is a struggle that has implications for surveillance. Data volumes have grown from about 10 million messages per day in some markets to more than 250 million as more orders are being entered and cancelled - a few seconds either side of an alert, they are probably covering about 20,000 events in a single instrument. High speed data feeds containing huge volumes from multi-asset sources need to be consolidated into a single comprehensive view to create a clear picture of normal market behaviour against which to proactively measure deviations.

**Big Data Buzz:** In surveillance, big data is not the next big challenge to conquer; it is a very current affair. Trade data is important, but to accurately understand market behaviour, systems need to analyse other types of external information including news, economic data such as interest rate changes, and even weather forecasts. Social feeds should also be examined - while difficult for exchanges to gather this information, market participants are already greatly investing and exchanges should look at how they can incorporate ‘big data’ into their surveillance programs.

**Beyond the surveillance system**

Today’s surveillance systems can provide value beyond traditional surveillance. The surveillance system provides the most accurate low-level store of market information held by an exchange, and advanced surveillance systems have flexible visualisations and data mining tools on top of this information, which can be leveraged by other teams.

Market operations teams can leverage the investment to become more proactive in boosting market control, like being able to reconstruct trading system events or to detect sudden surges/drops in order activity.

Front-office teams can leverage surveillance systems to see exactly where orders have been placed, when and at what price, greatly improving the service provided by the exchange to its members.

In conclusion, exchanges need an effective surveillance program in place to ensure market integrity, maintain investor confidence and grow the business. Technology is only one component. Strong rules and regulations, effective exchange processes and continual education and staff training are essential. As the market develops and matures, the surveillance program also must evolve to accommodate various players, different trading strategies and ever increasing volumes of data. MV

Proven in the world’s fastest and largest markets, SMARTS market surveillance and compliance solutions from Nasdaq are the industry benchmark for the global exchange, regulator and broker community. SMARTS Surveillance technology is installed in more than 50+ national exchanges and regulators and 100+ market participants across 65 markets.

**Contact Details:** SMARTS@nasdaq.com
Business.nasdaq.com/TECH
A tremendous new technology shift is about to impact the entire global exchange industry infrastructure. Heterogeneous computing platforms with field programmable gate array (FPGA) technology are implementing the next generation of exchanges. FPGA technologies provide ultra-low-latency, jitter-free operation, high throughput, and consume very little power as compared to CPUs and GPUs.

FPGAs are widely used for real-time networking and supercomputing application. FPGA technology uses fine-grain parallel hardware to accelerate exchanges so that financial transactions complete quickly and deterministically. Deterministic trading infrastructure satisfies government regulations for fair markets and the faster hardware scales to handle the traffic needed for global exchange connectivity and forecasted trading volume growth trends.

With roots at Stanford University, Algo-Logic Systems is the leading gateware developer for networking systems and builds multiple FPGA components for ultra-low-latency trading firms. It provides a broad range of highly scalable and reliable cores that run on off-the-shelf FPGA hardware platforms with high-speed Ethernet interfaces. These platforms have been deployed worldwide in co-located trading datacentres.

**Trade-to-Tick System**

Algo-Logic’s Trade-to-Tick (T2T) System receives orders from market participants and disseminates market data to subscribers. Within the T2T system, the order sequencer module assigns timestamps to orders and forwards the orders to the matching engines. The T2T System’s uniqueness lies in its ability to avoid delays and jitter caused by sequential software execution, bus transfers, memory copies, and cache misses that are present in standard software-only servers. The Trade-to-Tick path starts with Order Management Gateways that receive orders from market participants over the network. Sequencers benefit from hardware accelerated time-stamping of all the orders incoming from market participants. The Matching Engines use hardware accelerated in-memory databases, called a Key-Value Store (KVS), for maintaining open orders, aggregated order book building, sorting, as well as matching of outstanding orders based on price, timestamp, volume, and other priorities. Matched orders are distributed to clearing houses, order loggers,
The Trade-to-Tick System’s uniqueness lies in its ability to avoid delays and jitter caused by sequential software execution, bus transfers, memory copies, and cache misses that are present in standard software-only servers.

The sub-components within Algo-Logic’s Trade-to-Tick (T2T) System includes the Ultra-Low-Latency (ULL) PHY+MAC, Transmission Control Protocol (TCP) Endpoint, Full Order Book with aggregated in-memory matching engine, Protocol Parsers and Low Latency Library.

Ultra-Low-Latency PHY+MAC
To achieve the lowest Ethernet latency, Physical Level (PHY) and Media Access Control (MAC) functions are optimised for speed. Algo-Logic’s Ultra-Low-Latency (ULL) PHY+MAC receives and sends packets with a roundtrip latency measured in tens of nanoseconds. The ULL PHY+MAC design is compatible with multiple FPGA platforms that support high-speed Serializers and Deserializers (SERDES). The MAC interfaces to FPGA logic via standard 64-bit Avalon-ST or AXI4-Stream bus standards.

Ultra-Low-Latency TCP Endpoint
Algo-Logic’s TCP Endpoint maintains reliable connections for data flowing directly into FPGA hardware. It supports opening, maintaining, and closing TCP Connections. It has a latency of just 76.8 nanoseconds from the time that a packet arrives at the PHY/MAC to the time that the order processing logic can start to consume that data. By using multiple full duplex ports within a single FPGA, each device can handle hundreds of Gigabits/second of line-rate traffic. The implementation is portable between Altera and Xilinx FPGA devices and compatible with all widely deployed FPGA platforms from board vendors that include Nallatech, Terasic, Solarflare, Bittware, Alpha-Data, and PLDA.
Aggregated order-book with order matching algorithms

The order book implements the sorted matching function in an FPGA hardware device. In large exchanges, where scalability and redundancy are important, a distributed order book is used to handle large numbers of symbols and open orders. A fast, in-memory database is implemented using an FPGA-accelerated Key Value Store (KVS). Each Exact Match Search Engine (EMSE2) inside the KVS transacts 150 Million Transactions Per Second (MTPS). The KVS allows orders to be added, modified, and deleted with sub-microsecond latency.

Protocol Parsers

Algo-Logic’s protocol parsers extract data from the order flows. The parsers cover multiple markets for different exchanges across the globe. Parsers built in hardware include support for all major worldwide protocols, including FIX, OUCH for NASDAQ, XPRS for DirectEdge, BOE for BATS BZX, ArcaDirect for NYSE ARCA, Arrowhead for TSE, CME MDP3 for SBE, and Native Trading Gateway for LSE.

To facilitate new markets, an automated process generates a new parser, enabling rapid deployment of hardware in new markets and on-going support for changes in existing markets. The protocol parsers, combined with business logic, extract and process data as soon as fields of interest become available in the data.

Low Latency Library

The Low Latency Library provides a set of modular components that are integrated together to implement a complete trading system with the shortest possible time to market. These components include Internet Protocol (IP) packet decoding, TCP session processing, UDP datagram processing, and a register interface that allows the hardware to be easily controlled from a software API.

Because Algo-Logic provides such a wide library of pre-built ultra-low-latency infrastructure components that utilise user-friendly C/C++ APIs, minimal effort is required for integration into existing systems. These components are tested, validated, and deployed worldwide across multiple exchanges.

Conclusion

To compete effectively, exchanges are changing the way that they build their infrastructure. Rather than continue making further painful efforts to optimise legacy software systems, exchanges are turning to FPGAs to achieve higher throughputs with deterministic latencies. Global exchanges that recognise these advantages are the ones that will maintain their edge. As Charles Darwin said, “It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change.”

As Charles Darwin said, It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change.”

Contact Details: Algo-Logic Systems, Inc.
http://Algo-Logic.com
Email: solutions@Algo-Logic.com
Develxperts has been offering market operators a revolutionary approach to technology build since 2000. Originally a small software development company with expertise in broker technology, Devexperts changed that industry with platforms such as TD Ameritrade’s Thinkorswim. Our expertise derived from the experience of supporting brokerage operations 24x7 - hence developing the know-how necessary to also support exchanges’ technology requirements and provide superior servicing to their clients and members. Devexperts is currently participating in some of the exchange world’s most strategic projects, working with electronic communication networks (ECNs), alternative trading systems (ATSs) and exchanges to support secondary market operations via our unique, streamlined and cost-effective approach to delivering IT solutions.

The technology we use has evolved over the last 15 years – as has our approach to project management. Having developed the platforms of its early investors, Devexperts became savvy and built up its own internal development department, taking the knowledge accumulated over time and using it to develop the firm’s own line of products.

Devexperts offers a set of components, which can be put together like Lego blocks to build the foundation for a solution. With some additional customisation, that solution can be tailored to clients’ specifications. Those building blocks exist for each possible function - from a matching engine to a messaging bus. The system provides connectivity to internal and external information sources so that market data can be plugged in, normalised and then used within interoperable components - fully integrated, yet flexible.

Building an exchange is a complex project, which makes the componentised approach to technology invaluable. The building blocks can be combined in different configurations to build different solutions, hence bringing flexibility and speed to the process of customising solutions to any given challenge – yet based on proven technology. While exchanges will typically look for assistance with their back office and services, Devexperts can provide a full front-to-back service offering, with a full set of front-end components - including exchange terminals supporting different asset classes and displaying all of the information that the exchange is giving out. It uses Java-based desktop technology giving it the capacity to run across operating systems, Mac or Windows, without any impact on experience or performance.

The HTML 5-enabled system offers a comprehensive set of desk-top terminals, mobile terminals and HTML 5 web terminals across smartphones, tablets and PCs.

Data: Capturing the market

Two issues dominate data provision and analysis for market operators today: the drive towards greater transparency for all market participants, and the pressure to improve market supervision. They both require a higher level of information granularity than previously needed, with a greater capacity to query and analyse captured data.

To support market operators, Devexperts’ platforms allow them to look at real-time market data, historical data and reference data across symbols or in an aggregated view, according to the trading venue’s need – hence addressing the requirements for market transparency and reporting.

To deliver greater market transparency, traceability and audit, Devexperts provides a data solution with a capacity on par with the Consolidated Audit Trail (developed by the Securities and Exchanges Commission, in the US) and with a five-year proven record. The solution currently captures data from all of the US exchanges in order to provide a traceable log of what happened at any given point in the market.

Data capture, storage and retrieval is available right out of the box - but it can have its pre-defined functionality extended through customisation. The model Devexperts uses is data agnostic since any data provided to the platform is normalised. As such, the system is able to handle information from multiple sources in multiple formats and protocols. Market data could be well defined and highly structured, but it can also be more open, so that the set of data captured is expansible - hence allowing for future requirement changes. For example, it might be expanded to increase the granularity of data around a specific characteristic, or it might take in new details such as the requirements for trader identification under the Markets in Financial Instruments Regulation (MiFIR).

The market snapshot that the system provides is accessible for the exchange’s internal processes but it can also be made accessible to clients in order to dramatically increase transparency.

Since both the marketplace and its members have the same information available to them, there is an ‘official history’ of what has truly happened in the market removing doubt around market events and
supporting smaller firms in their capacity to keep up to increasing complexity and volumes of data.

Given the current concern around volatility events, this ability is highly valuable. There would be no question about what happened in a ‘flash crash’ similar to the event of May 2010; to find out what happened the exchange’s audit team would simply dig in and see who conducted what activities. This functionality could also be utilised for a market abuse inquiry or meeting any other regulatory requirement.

Most of the historical and real-time aggregated data are accessible through a single set of services and application programming interface (API). From an operational point of view this means there is no need to build different interfaces to accept different types of data. The user can just switch a couple of parameters to change from historical to real-time and between aggregated or reference data, making it a simple model to use. Data types that are typically used by an exchange can be pushed into the Devexperts normalised structure, making life easier for those who need to develop technology solutions to run on top of the exchange platform. Be it client systems or internal systems, processing data or storing data, Devexperts has that technology.

Our technology is built entirely in-house – as such, there are no third-party production sites. For historical data storage systems there are no third party products and therefore no third-party license costs in Devexperts technology projects. It is all a single package which the firm completely owns and offers complete support for.

**Market data**

To manage market data Devexperts offers a low-latency messaging platform that handles real-time data and utilises an internal system to distribute that data. This system can also function as a standard external protocol through the use of a separate API, - allowing customers to connect and get the data with very little code development. The platform supports Java, C++ and most web services. Devexperts can deliver an exchange project with a full set of protocols right out of the box with compatibility with every major programming language, and every major technology. It supports standard distribution mechanisms such as Unicast and Multicast; it can also be converted to use Financial Information eXchange (FIX) protocol if a client prefers to deliver information via FIX messages. Alternatively this can be a subscription-based service via which clients can receive a subset of symbols from the exchange’s data feed. Depending on configuration, the system can deliver bulk feeds or subscription-based live feeds.

**Historical data**

Devexperts has a unique historical data technology. The firm has been storing Level 1 and Level 2 tick data since 1 January 2010 from every American exchange across all asset classes: equities, options, futures, indexes and FX. The amount of information captures is extensive e.g. quotes in the US options market, which due to the number of instruments listed and the amount of quotes given across US exchanges is among the largest feeds in the world. This data is instantly accessible in real-time and it is also available from a historical database – retrievable from the moment subsequent to capture all the way back to the start of the data capturing period in January 2010.

As an out-of-the-box feature, the Devexperts technology makes data available from any period in the reporting history in a matter of seconds. Typically this is sold as a service, providing clients with access to the cloud platform, with data and a set of client-facing services, APIs and subscription services. One of the main features it offers is a market replay from the cloud, which is akin to a time machine for the market. Either via API or from the web it can replay specific moments in time for symbols covered by the market data subscription at the users’ preferred speed. Whether yesterday or one year ago the time to provide a particular replay takes approximately five seconds or less, depending on the nature request.

While the replay service streams directly from the cloud, a second scenario is possible using a Devexperts query engine. The web service query engines lets the exchange extract the data for a specific symbol, experience, or event based on user-defined parameters, upon which it extracts the file or streams it.

**Supermodel**

Rather than use a relational data model, which is used in traditional SQL databases and data warehouses, Devexperts has developed its own engine for historical data. It is a database in the broader sense; a storage engine for unstructured data which is then structured to provide quantitative analysis. This big
data engine is able to run queries against the data and process it efficiently in large amounts.

To run high-speed analysis of data, the platform uses an on-demand, big-data processing model. It integrates with the Spark Framework, which lets the user run queries across clusters of processors, provided on-demand, creating a parallel-processing capability without any need for owned infrastructure. At present, all the historical data that Devexperts holds is compressed at a ratio of 1 to 10, leaving the firm managing about 200 terabytes of compressed data which represents five-and-a-half years of data from the American market. Uncompressed it would be nearly two petabytes.

With this technology users are able to address and query data, while providing constant query times across all time periods. As a consequence, when the user is extracting the data, or processing data with an on-demand cluster or just replaying data, Devexperts provides very fast response times. Those response times are not impacted by the depth of history that a query accounts for, or by the type or number of symbols that are being queried.

Since the platform covers all aspects of real-time and historical data, it fully underpins the provision of data services to exchange members as well as supporting the internal processes of an exchange. This is done by providing access to data as and when needed, regardless of the formats that might be used.

**Balance of strength and flexibility**

Flexibility is built into the system, and wedded to the reliability of the platform itself. As such, the exchange gains additional resilience in tough circumstances without needing to adopt the kind of rigidity that may limit the platform as regulatory demands change in the future.

The structure of storage is defined along with the data, so the system will apply any changes to the data model across the entire platform, redefining or extending the data scheme and extending those changes into data storage. Although feed handlers for specific exchanges are built in proprietary code the messaging engine is open source – and is currently in its fifth generation. Access to the storage engine varies according to the licence type.

In order to tailor this model to suit clients across geographies and instruments, Devexperts offers a variety of approaches to the development model. Teams of coders are accessible to clients in order to develop systems specifically for individual customer’s needs. Alternatively – and perhaps more suitably for new exchanges, a lighter approach allows a shared team to be created between the client and Devexperts in order to leverage our many ready-made components with the client’s demands. This process can reduce IT expenses and enable faster times-to-market – while still delivering very specific features. The use of shared teams is highly recommended in environments where ready-made components can enable the very fast deployment of custom-built solutions.

The market operator is hence able to decide whether it needs a dedicated technical capability or a lighter support model with only occasional development of product - or maybe a monthly development cycle.

Above all, Devexperts offers flexibility – we encourage different approaches to different requirements. MV