COMPONENTS
The Panopticon platform includes:

- **Designer (Windows)**
  - Data Connection & Drag & Drop Analytical Dashboard Design

- **Web Client**
  - Interactive Visual Analysis through Analytical HTML Dashboards

- **Server (Windows / Linux)**
  - Powering the Web client and also provides a real time caching data conduit, report generator, and enterprise integration.

The web client is certified against Chrome, IE10+, and Safari on iOS.

USAGE
Real Time streaming analytical dashboards are used not just for live monitoring, but for intra-day and historic analysis. Specifically to identify the anomaly as it occurs, investigate its evolution, and then analyze history to gain insights on how to act. Panopticon reduces delays caused by human analysis and decision making to avoid financial impact. In Capital Markets that spans the entire trading lifecycle, from analysis of e-trading flow, to analysis of risk in open positions, to the surveillance of that flow investigating market abuse. In Manufacturing, Energy and Utilities that includes predictive maintenance plus capacity and flow analysis; this time monitoring sensor data, rather than market data, orders and executions. Panopticon reduces the time necessary to understand, gain insights and then act on operational datasets.
Events can occur during the blink of the eye, and can only effectively be monitored electronically via pre-defined rules engines. However these rules are based on correlation with historic performance, so when events diverge from this, human analysis is required. The ability to receive, analyze, understand, and act with minimal latency has a direct and immediate financial impact. Providing the ability to generate and then visualize the latest operational state, whether push subscribing to a feed, or automatic pull data requests from a repository delivers on this objective. Typically rather than looking at raw data, metrics are displayed comparing actual performance against expected performance, highlighting specific anomalies, whether visually or through desktop alerting. Panopticon has the ability to jump from real time views to intraday and historic views, allowing investigation of the anomaly to continue, which may result in both informed action, and updates to the external rules engine to cater for these new market events.

The investigation of anomalies is enhanced with both viewing, and interacting with the historic events. Being able to view the divergence evolve across time, and then resample down to every transaction, allows the impact and evolution of events to be quickly investigated. Additionally, rather than focusing on the anomaly specifically, playing back through history of all flow, whether long term across days or for a specific day down to the millisecond, improves understanding of the macro view on what actually occurred, highlighting opportunities and risks in the larger dataset, which can then be quickly investigated.

The system acts as a time machine, allowing you to identify time based trends, clusters, correlations, and abnormalities, and then to focus on these, to step through time around them, resampling as necessary to drill down to the individual events.

To understand, gain insight and take informed action takes time. Time that can have a direct financial impact the longer the latency is from data being received to informed effective action. However due to the sheer volume of data which itself is continually changing, it is impossible in most cases to read every individual number. Summarizing the data in charts to tables, reduces the data volume problem, but tends to show you the expected metrics that don’t change, or if they do change, no understanding of what’s driving them. Instead we need to see more data, and we’re back to being overwhelmed. Panopticon uses pre-attentive techniques, to visualize data in high density displays appropriate to the questions being asked, whether peer comparison, trending, clustering, correlations, relationships or outlier detection.

Panopticon provides high density visualizations of operational metrics that rapidly highlight divergence from expected performance. Rather than spending minutes reviewing tabular reports, or viewing high level BI reports and missing crucial details, the abnormality is instantly highlighted.

OEM PARTNERS
In many scenarios, data items do not reside in isolation, but instead fit into defined hierarchies, such as business, instrument, venue and customer. Efficient identification of peer group outliers and performance clustering provides the starting point for investigation, insight and informed decision making. To achieve this requires both the aggregation of selected metrics, the calculation against expected returns, and the quick identification of anomalies through high density, hierarchical displays. Panopticon provides each of these, offering for example a set of alternative hierarchical displays, where the end user can dynamically select the detail level, and rearrange the hierarchy as required to identify problem performance.

In certain scenarios the data will be too complex to aggregate locally, and instead will need to be calculated in a grid or real time cube. This is typically the case when working with risk exposures and sensitivities. Panopticon has the ability to swap to supporting externally provided aggregates, allowing analysis of datasets that cannot physically fit into the power of a single machine. Data can then be visualized hierarchically, so that diverges from expected returns within the tree structure (due to divergences from expected correlations) can be highlighted even if the top and leaf values are as expected.

Traditional bespoke user interfaces are long term and expensive development projects. Whether in Java, WPF or HTML through D3, the illusion is that they are simple, and sometimes due to the open source libraries utilized “free”. In reality the implementation of real time analytical dashboards that are not just read-only reports, but interactive allowing immediate action to be taken from gained insights is complex, and requires significant skilled resources.

The ability to design and deploy displays that support review of the big picture, identification of unusual behavior and subsequent investigation without coding dramatically reduces delivery risk, and delivery time. Instead user interfaces can be iterated by and with end business users that understand their datasets and their objectives. Delivery to the web browser is via one click, so the focus is on the data and the insights that can be gained, rather than costly development projects.

These user interfaces become more powerful when they are hooked into underlying systems, and tick databases, so that scenarios can be replayed, parameters input or updated and results compared.

Order & Execution data is typically captured, stored and analyzed within a time series or tick database. With data volumes in the Terabytes, and the specific time series analysis required, whether joins, transformations, calculations, aggregation and conflation, most of our customers have adopted tick databases such as OneMarketData’s OneTick, and Kx’s kdb+.

To optimally utilize these analytical engines, it is not enough to connect, and extract data, for example via rudimentary ODBC interfaces. Instead the tool’s power is released through native interfaces and close interaction. Copying data to a separate repository, removes the special capabilities of these tools and introduces unnecessary latency, delay and risk. However, native access, both request / response for intra-day and historic, and subscriptions for live analytical streams moves the end consumer of analytical dashboards closer to the underlying tick engine.

Panopticon’s connectivity is continuously being enhanced to extract maximum value, and reflect the requirements of both etrading teams, and the associated compliance surveillance teams, which typically include how to join executions and market data, how to highlight order book imbalances, how to aggregate positions and executions, how to calculate metrics, and how to dynamically conflate to the appropriate time granularity.
The requirements for Big Data analytical tools have moved beyond the batch environment available through Hadoop, and Map Reduce, to interactive analysis, where timing is critical. As a consequence the big data technology stack has moved on to reflect this new reality with Cassandra for fast storage and querying, Spark for fast analytical querying including aggregation, Kafka for enterprise messaging and Panopticon for real time streaming analytics as well as intra-day and historic analysis.

For maximum effectiveness, analytics views must include real-time data when possible. When there are significant financial penalties for being late, it isn’t practical to wait for a report to be refreshed, or manually keep clicking “refresh”. Instead the platform needs to subscribe to the live metrics, typically showing current and cumulative flow, and do so in such a way to not be overwhelmed.

Subscriptions to message buses, CEP Engines and “Live” databases, requires the data be subscribed against, throttled to minimize CPU and memory consumption, and output to analytical displays maintaining state for the latest data, or delivering scrolling time windows. When push subscriptions are not available, data has to be pulled regularly from underlying repositories, whether databases, cubes, or restful web services. Even in these cases analytical displays update to always show the latest state of the underlying data environment.

As analytical requirements become more statistical, moving from the analyst to the data scientist, the requirements for predictive or advanced analytics increases. For most existing customers this has meant either utilizing R or Python, with their associated modules / libraries.

Rather than duplicating predictive capabilities inside the system designer for high performance visual analysis, we extend to support R and Python through either data transforms as part of the data retrieval pipeline, or additional data sources. For example to perform cluster identification using K Means, data can be retrieved through any of the available sources supported, auto-generated as a data frame, pushed to Python or R with an associated script to perform analysis such as K Means clustering, and returning a resulting data frame that can then be displayed. Alternatively if the data and analysts already exists in either Python or R, they can act as data sources, and be automatically polled for updated data. Data input allows the script to be dynamically updated, allowing the end user to both control the data source, and the analytical execution.

To support mission critical operations any system has to both deliver on functionality, and then deliver on throughput, performance, scaling, and fit into an organization’s technology environment, both for initial deployment, promotion between environments, and ongoing support. For most customers this means deployment on Linux close to their data infrastructure, security integration, audit and operational logging, plus jmx support for operational performance statistics.

Internet facing deployments typically add complexity with requirements for splitting the server across a DMZ and internal network segment, while binding into the single sign on mechanism of the parent web portal. In addition to providing secure access to analytical dashboards through a web client, the server also needs to produce offline PDF reporting, where reports can be scheduled for production and download, or emailed to defined distribution lists.