VANTIQ –

The Connected Industries Integration Platform

Paul Butterworth, CTO

April 2019

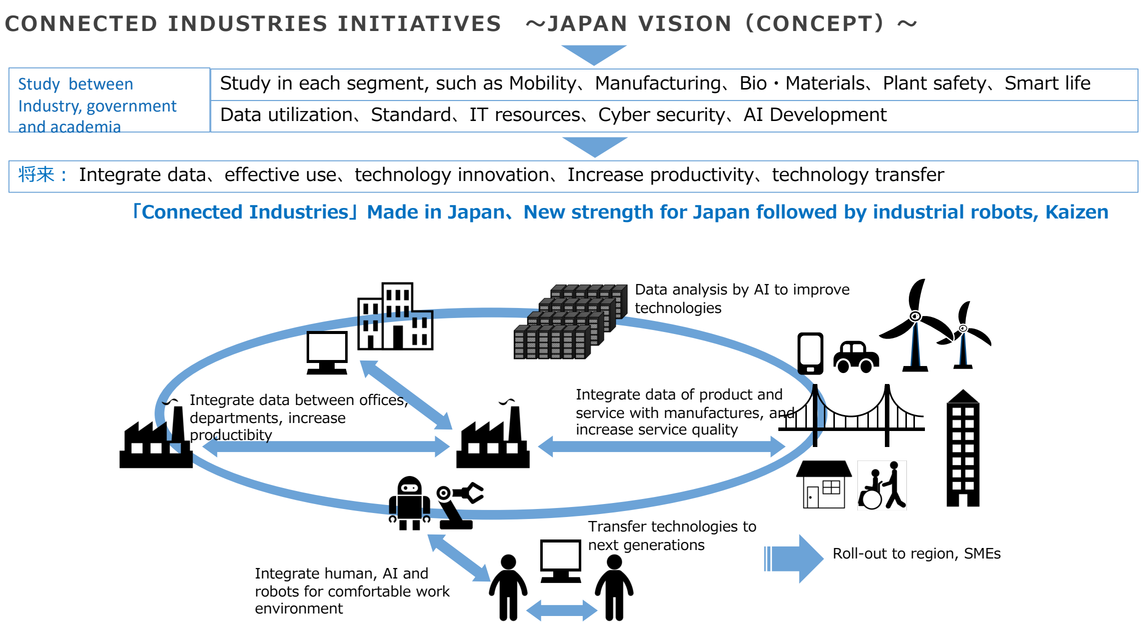
Executive Summary

The concept of the 'Connected Industry' involves connecting machines and people across boundaries such as geographies and time. Maximizing the potential of connected industries demands platforms that:

- Integrate technologies such as AI/ML, remote sensing and real-time analytics
- Connect people and things
- Operate in real time
- Operate across distributed environments

METI, the Japanese Ministry of Economy, Trade and Industry has been working on developing a vision for a so-called 'Connected Industry initiative', a "new concept framework in which industries will create new added value and the solutions to various problems in society through connectedness of various facets of modern life, including humans (including our roles as consumers and suppliers), machines, systems, companies." The purpose of this document is to describe how VANTIQ is the perfect platform for executing against similar initiatives worldwide.

This initiative is depicted in the following METI diagram:



Connectivity among machines and people using advanced technologies such as Machine Learning and Remote Sensing must be real time in order to enable the real-time business of the future. Such real-time connectivity is best enabled by a real-time, event-driven integration platform that supplements traditional non-real-time data integration platforms as depicted in the diagram below.

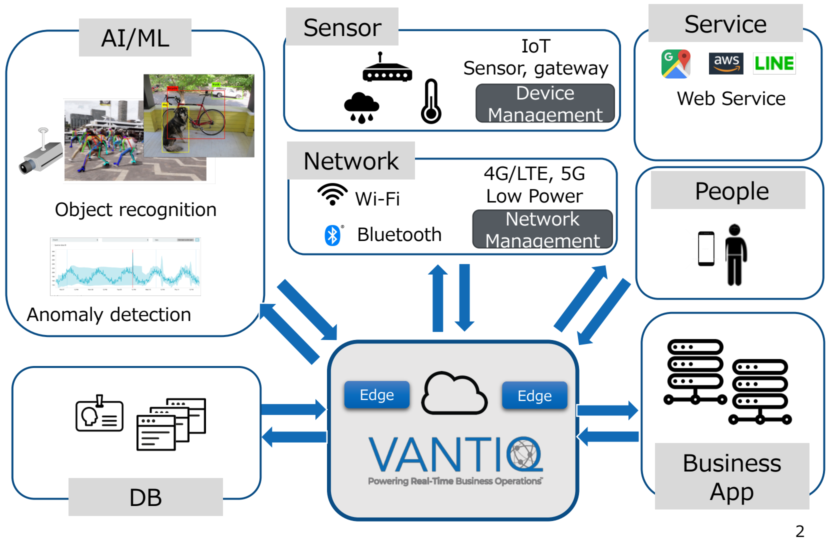
Integration Platform for Machines, People, Sensors, AI

Maximizing business value, integrating sensors, AI/ML, web services, legacy systems, people, apps, databases, etc.

Combining and analyzing the dynamic and static data in real-time using advanced event processing platform.

Dynamic data
Asynchronous, streaming data from sensor, AI/ML, people (smartphone), Business Apps.

Static data
Database and legacy systems



VANTIQ represents the state of the art in real-time, event-driven integration providing the high levels of functionality required for advanced event-driven integration in a distributed environment operating across the cloud, on premise and at the edge. VANTIQ provides the scalability, reliability and manageability demanded by complex, global systems. In addition, VANTIQ offers event-driven collaboration tools enabling a more effective model for integrating advanced automated systems with the people that use and manage those systems.

VANTIQ offers a state-of-the-art platform for satisfying the integration needs of Connected Industry.

Connected Industry and Integration

The vision of Connected Industry requires integrating all elements of industry into a connected whole. The elements being connected include:

- Technologies
    - Sensors
    - Artificial intelligence and machine learning
    - Robotics
    - Advanced, real-time analytics
    - Legacy systems and databases
    - 4G, 5G and low power networking
- People
- Companies
- Industries
- Countries

Connected Industry will require this integration to operate in a technical environment that is:

- Distributed – connecting factories, offices, commercial outlets and homes distributed throughout the world.
- Real-time –connected industry must operate in real-time to reach its full potential in supporting real-time business.
- Heterogeneous – both in computing and networking.

These requirements motivate the use of a platform that can support all aspects of the integration problem:

- Data acquisition for collecting information from sensors and machines, environmental data from public sources and social media, commercial data form legacy applications and databases, and inputs from the people that use and operate the systems.
- Analysis of the data in real-time to determine what actions are justified. Analytics may be rule-based or statistical and may involve the use of various forms of artificial intelligence, machine learning and recommender systems.
- React in real-time either by direct control of actuators, control through the supervisory systems that manage the equipment or collaboratively by efficiently integrating people with the machines and information systems that are part of the Connected Industry.

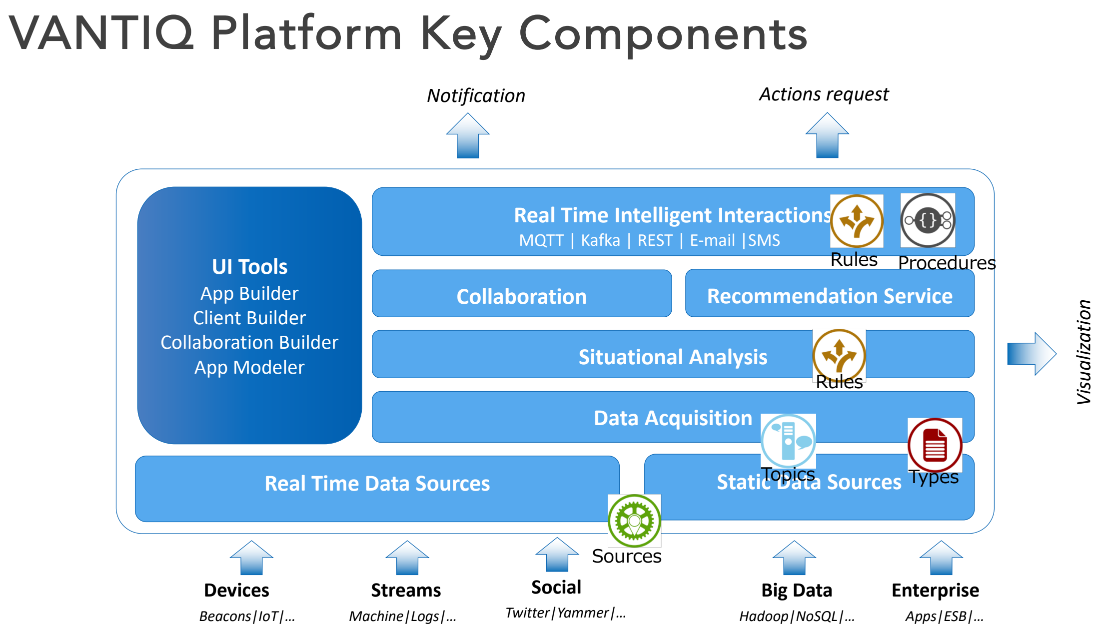
Finally, the process of integrating the Connected Industry elements must be highly efficient and agile given the extent of the initiative, the continuous advances in the technologies involved and the ongoing evolution of real-time business.

The most efficient approach to achieving these goals from the technical perspective is to use emerging real-time, event-based integration architectures realized on a high-productivity, agile, event-driven platform.

VANTIQ

VANTIQ offers a state-of-the-art platform supporting both event-driven applications and event-driven integration. VANTIQ tools offer high productivity reducing time to market and lower costs for new initiatives. The VANTIQ runtime platform offers the mission critical capabilities necessary to support large, distributed event-driven systems.

The key components of the VANTIQ platform are depicted below:



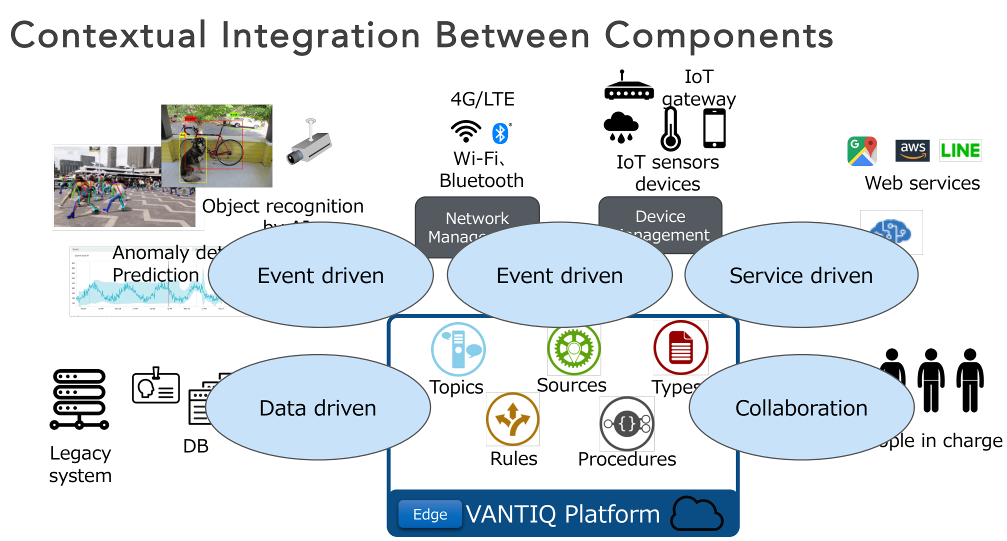
VANTIQ Event-Driven Integration

VANTIQ offers a full range of integration architectures (See *Appendix A: Integration Architectures Overview* for a definition of each type of integration) including:

- Event-driven integration
- Data-driven integration
- Service-driven integration

VANTIQ offers this suite of integration architectures because there are integration activities in connected industry for which each architecture is best suited. However, VANTIQ focuses its efforts on event-driven integration architecture that serves as the foundation of real-time connected industry.

The optimal use of the various integration architectures is summarized in the diagram below:



Integrate dynamic events to make optimal decisions

Gartner identifies the [Advanced Event Broker (AEB)] as a key enabling technology. Essentially, this is any technology that transports events from their source to their destination. However, most AEBs focus exclusively on the transport of events between the publishers and subscribers. This is certainly necessary functionality but insufficient to support the complex and varied integration needs of Connected Industry. For example, VANTIQ event-driven integration offers the following capabilities that are required to quickly and efficiently integrate machines, people, companies and countries using technologies such as IoT, artificial intelligence, machine learning and real-time analytics:

- Transformation – converting events to a form that is simpler to consume by subscribers. For example, a temperature sensor may transmit a voltage that is then transformed into a temperature for consumption by applications. Farther downstream this temperature may need to be converted from Celsius to Fahrenheit for simpler consumption by specific subscribers.
- Filtering – not all events are relevant and allowing them to flow downstream wastes time and resources. Filtering suppresses irrelevant events. For example, if the temperature sensor above produces an event once a second containing the ambient temperature, it is unlikely the temperature is changing every second. The sequence of events can be filtered so that only temperature change events are allowed to flow downstream to subscribers.
- Contextualization – the payload associated with the event may be augmented by attaching additional information to the event. For example, a temperature change event may be augmented with the history of temperatures reported by this sensor or the augmentation may require submitting the event to a machine learning system that has been trained to detect anomalies in the pattern of temperatures reported. The anomaly may then be reported as

another event, effectively combining contextualization and filtering in the processing of the event.
- Correlation – events sourced from multiple event streams may need to be correlated in order to be effectively processed. For example, a safety system may be producing an event stream of biometric information for an employee while a separate event stream is reporting the location of the employee. These two streams must be correlated in time to produce a new event stream of employees whose biometric data indicate a health problem correlated with the location at which the health problem is occurring.
- Streaming Analytics – analyzing event streams in real-time is critical to making real-time decisions. VANTIQ event processing can apply real-time analytics directly to the event streams as they are received. Statistical methods as well as AI/ML approaches to analysis may be applied.
- Time series analysis - VANTIQ's built-in real-time analytics automatically and efficiently manages time series data to produce statistical measures such as moving averages and rates of change. This solution eliminates the need to store intermediate series values in a database dramatically improving real-time response and processing efficiency.
- Custom Processing – a common requirement is to apply specialized business logic or analytics to the event stream. VANTIQ event-driven integration offers a number of additional integration facilities for incorporating such specialized logic into the event-driven integration flows.
- Collaboration – people are a critical element in Connected Industry and VANTIQ offers highly optimized tools and platforms for integrating people with systems in the most effective manner possible. Collaboration will be discussed in more detail in a subsequent section of this document.

Teams responsible for Integration on a global scale require a catalog of available events in order to effectively construct integrated systems. VANTIQ offers powerful Event Catalogs hosting the inventory of available events published throughout the Connected Industry ecosystem. The teams responsible for constructing new systems within Connected Industry can search these catalogs for relevant events that are available throughout the distributed environment and then use the associated metadata to quickly incorporate the selected events into the systems they are constructing. The VANTIQ Event Catalogs can be deployed on a global basis or on a more restrictive basis for events specific to an industry, geography or company. This provides a level of visibility and security commensurate with the intended uses of each event catalog.

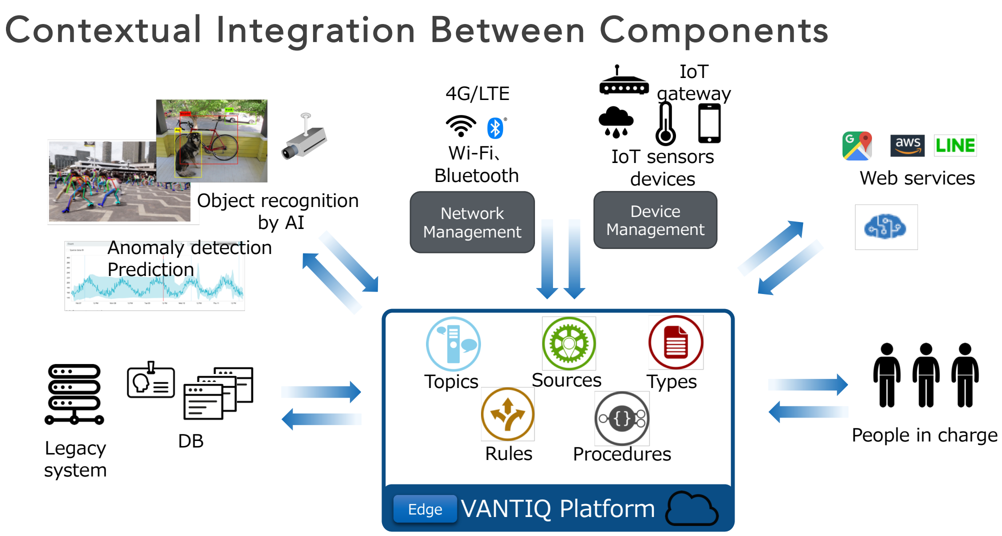
In addition to these functional capabilities, VANTIQ operates with a full set of mission critical quality characteristics:

- Scalability
- High availability, resilience and fault tolerance
- Security
- Audit trails
- Operational analytics
- Optimization

- Tracking and billing

Connected Industry Integration Scenarios

Examples scenarios illustrate VANTIQ's unique approach to event-driven integration for Connected Industry.



Integrate dynamic events to make optimal decisions

Event-driven Integration for AI and Machine Learning

Connected Industry will make extensive use of emerging machine learning technologies. VANTIQ's position is that machine learning and other artificial intelligence technologies are still emerging and the best approach to AI and machine learning has yet to emerge. A number of machine learning technologies are available each of which does an effective job of decision making in specific domains. This state of affairs has led VANTIQ to support the integration of a wide variety of artificial intelligence and machine learning systems into the real-time, event-driven systems produced in VANTIQ using multiple integration architectures.

Once integrated, the machine learning system analyzes real-time event streams as follows:

- Identify event streams that should be analyzed by a machine learning system.
- Direct the event streams to the machine learning system which responds to each event with an analysis.
- The analysis is published as a separate series of events that can be used by any number of applications in their decision making.

For example, in an actual VANTIQ system, video images from a set of cameras monitoring streets are fed to a Tensorflow model. The Tensorflow model identifies cars, trucks and pedestrians in the street view. One application uses this information to identify problematic traffic conditions, using the VANTIQ collaboration system to integrate the relevant managers into the response to the traffic condition. A second application uses this information to report on potential safety issues for pedestrians, in this case using collaboration to integrate the relevant authorities into the response to the event. Similar approaches can be applied to operational and safety systems in Connected Industry.

In another example, integration with cloud-based machine learning services is utilized to convert natural language into directives to a field service management application allowing the field service personnel to perform their duties using hands-free interactions with their support system.

Using multiple integration architectures, VANTIQ integrates the optimal machine learning system into each connected industry solution to provide the most accurate analysis available.

Event-driven Integration, IoT and Connected Industry

IoT is a key enabler of Connected Industry. The previous section on AI/ML outlined how VANTIQ integrates video streams with the intelligent analytics models that interpret those streams. Video streams can be considered one class of IoT device. VANTIQ integrates event streams produced by any type of sensor and can transmit control commands to any type of actuator by using the standard connectors offered with the VANTIQ platform. If standard connectors are inadequate VANTIQ provides an open source toolkit for constructing additional connectors.

As noted in a previous section of this document, simply capturing the data does not solve the problem. The data must be transformed, contextualized and analyzed in real-time in order to take full advantage of the data throughout Connected Industry. The VANTIQ platform offers both platform capabilities and tools that simplify the construction of such “intelligent” integrations and operate them in an efficient manner.

Finally, sensors may produce large volumes of events. Large volumes of events should always be processed at the edge in order to reduce latency and data transmission costs. For example, safety applications that monitor the state of operating equipment must respond immediately when an anomaly is detected to reduce the possibility of injuring nearby operators or of serious damage to the equipment. Sending the events to the cloud, analyzing and then responding back to the local operating environment is too slow for such applications involving tens, hundreds or, worst case, thousands of milliseconds of delay for data transmission. In contrast, appropriately configured edge processing can respond in just a few milliseconds. VANTIQ’s distributed architecture allows all event-driven processing, including analytics, to operate at the edge in a completely transparent fashion. In addition, VANTIQ offers a unique “partitioning” capability to allow the integration team to dynamically and declaratively move event-driven processing from the cloud to the edge or back in order to quickly establish the most efficient allocation of processing between the cloud and the edge.

Event-driven Integration, 3rd party Systems and Services

Connected Industry incorporates many existing applications and systems into the global solution. VANTIQ offers a number of standard connectors for integrating with existing systems and databases. VANTIQ also offers an open source toolkit for building additional connectors to existing systems and database that have proprietary APIs.

Modern application architectures exploit services and micro-services to increase agility and reuse. Service-based architectures will be heavily used with Connected Industry initiatives.
VANTIQ offers a complete set of built-in capabilities for integrating services and micro-services into Connected Industry applications. In addition, micro-services can be constructed within VANTIQ and deployed in a serverless fashion resulting in the benefits of both micro-services and serverless deployment.

Event-driven Integration and People

Connected Industry requires integration not just between machines and systems but between machines, systems and people. People, working collaboratively with machines and with each other, play a critical role in optimizing decision making within Connected Industry. VANTIQ's collaboration system offers a unique approach, maximizing the opportunity for people and systems to collaborate in an optimal manner.

Collaboration Examples

Some collaboration examples from existing systems are instructive:

> Companies such as Uber, Amazon, Deere and Tesla are creating new business models using software as an enabler. They take maximum advantage of real-time human-machine collaboration, often in unforeseen ways. Some examples:
>
> - Links in a supply chain being managed by humans in real-time with the system making recommendations that take specific customer needs into account
> - Floor managers in a casino collaborating with an application that combines guidance from ML (Machine Learning) algorithms with data from systems of record to create special offers for high-rollers to keep them engaged
> - Drones delivering packages to a delivery vehicle in transit and the car only has to slow down or pull over as the package is delivered
> - Drones and robots collaborating with firemen to put out a fire
> - After an earthquake at a construction site, the systems and the operations personnel collaborate to make numerous critical decisions

Faster response by humans is critical in many situations. Getting doctors, nurses, firemen, and police involved quicker with better situational awareness can save lives. As an example, a study in India indicated that 40% of children after open heart surgery died with the primary reason being that it took doctors on average 90 minutes to get to a child in distress because of ineffective communication. In rugged environments such as offshore oil rigs or underground mines, the benefits of quick action when problems occur are evident.

VANTIQ Collaboration Architecture

Integrating situationally aware software with decision-capable humans by creating sophisticated collaborations between humans and machines is very difficult. A Netflix or Uber, each with hundreds or thousands of very talented software professionals can do it over long-enough periods of time. However, addressing the challenge of creating rich collaborations in a timely manner with fewer resources requires building the collaborations using higher levels abstractions.

VANTIQ has incorporated advanced collaboration facilities consisting of a unique set of collaboration primitives, called collaboration patterns, as abstract components. Collaborations are created with a visual design tool. They inherit all the mission-critical capabilities VANTIQ provides. After initial deployment, VANTIQ collaborations can be dynamically modified as the system requirements change and deployed with no downtime. Within a VANTIQ application, powerful collaborations can be initiated at any time.

Event-driven Integration and Distribution

Distributed Business Applications

Many real-time, event-driven connected industry applications are naturally distributed. In manufacturing environments, Programmable Logic Controllers (PLCs) communicate with area controllers and edge nodes that forward the data to more centralized IT systems. In consumer environments, data may be collected from numerous position sensors, processed locally into logical locations on which immediate automation decisions are made and forwarded to remote systems that optimize the experience for the consumer. Such a wide variety of distributed applications require support for an equally broad set of distributed topologies ranging from devices directly reporting to a central site, to hierarchically structured automation systems, to federated peers collaborating to improve coordination across organizations or businesses.

The simplest architectures are sensors reporting to a central site. Many examples of such systems exist today. A system collecting sensor data from a mobile phone and reporting that data to a cloud service represents a common example of a centralized architecture. These simple *star* architectures represent the bulk of the existing event-driven business applications as they are the easiest to understand and build.

Hierarchical architectures contain additional levels of processing and connectivity. Hierarchical systems are more complex and mimic many existing physical and organizational structures. For example, an industrial IoT system that consists of sensors reporting to local controllers that report to plant-wide controllers that report to divisional headquarters that report to corporate headquarters represents a tree topology. These systems provide both centralized and decentralized monitoring and control. Such systems are more responsive in real-time or near real-time situations. For example, it would be impractical to control factory equipment in real time by collecting the data, shipping it to corporate HQ and having corporate HQ systems determine the next action for the machine. It is far more effective to do such an analysis on the local controller and simply report the situation and the action taken to the plant-wide controllers and, subsequently, to regional and corporate HQ. Faster response times, improved availability and local control make the distribution of the situational evaluation, collaborative decision making and response processing across the hierarchical topology more efficient than moving everything to HQ and making all decisions in a centralized fashion. Another classic example of hierarchical, real-time, event-driven business applications is the use of edge nodes to act as local processors for a collection of sensors and control points with the edge nodes then interacting with more centralized systems.

The most sophisticated distributed real-time, event-driven business applications are peer-to-peer systems where peers are managed by separate organizations. For example, in an electrical demand-response system, the overall system consists of sensors managed by power utilities and sensors managed by utility customers while control of the system is distributed across the utility and its customers. To provide real-time demand-response, the utility system and the customer systems must collaborate. This is accomplished by each system making local decisions and transmitting both the local situation and the local decisions to the other party and then agreeing to modify their real-time behavior based on feedback from each other.

VANTIQ Distributed Architecture

VANTIQ supports a general model of distributed and federated topologies. A distributed VANTIQ system consists of two or more nodes with each node representing a VANTIQ installation. A VANTIQ installation can contain a single service instance or a cluster of service instances. The VANTIQ installations are assembled into a distributed topology when an installation declares at least one "peer" node with which it desires to exchange messages.

VANTIQ installations, by default, are independently managed. A node, A, declaring another node, B, as a peer MUST have credentials to access node B. Thus, the system is naturally federated since a node may only exchange messages with another node if it has been granted sufficient rights to perform the desired operation on the peer node. Peering is symmetric. If node B wishes to exchange messages with Node A, Node B must provision Node A as a peer and have sufficient rights to access node A.

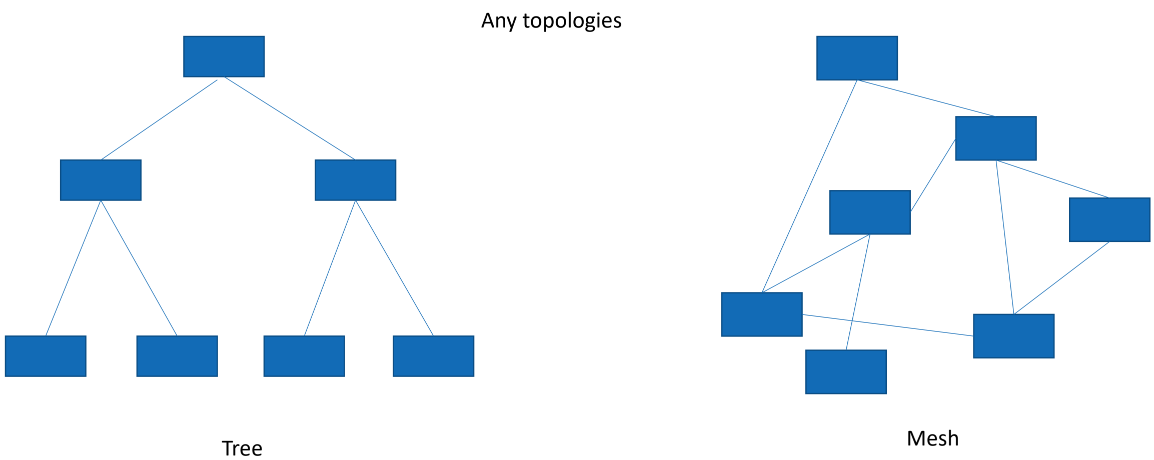
Since the peering relationships can be defined between any two nodes, VANTIQ can support any distributed topology. Also, the topologies are implicitly federated since authentication and authorization are independently managed at each node.

VANTIQ anticipates that initial usage patterns will favor topologies in which all nodes in the distributed system are managed by a single authority. Such systems are typically organized into star and tree topologies:

- Star – consists of a single parent node with an arbitrary number of child nodes.
- Tree – consists of a root node with an arbitrary number of child nodes where each child node may act as a parent for an arbitrary number of child nodes.

As the deployed system becomes more collaborative, federated peer-to-peer networks will be constructed. In such a network topology, any node may peer with any other node leading to a general graph structure representing the connections among the nodes. The mesh network model tends to be the most complex since cycles in the graph are possible and the cycles must be handled by any functions that operate on more than one node in the graph. Also, because each VANTIQ node represents an independent system that may require separate credentials the systems naturally generalize to federations among collaborating organizations.

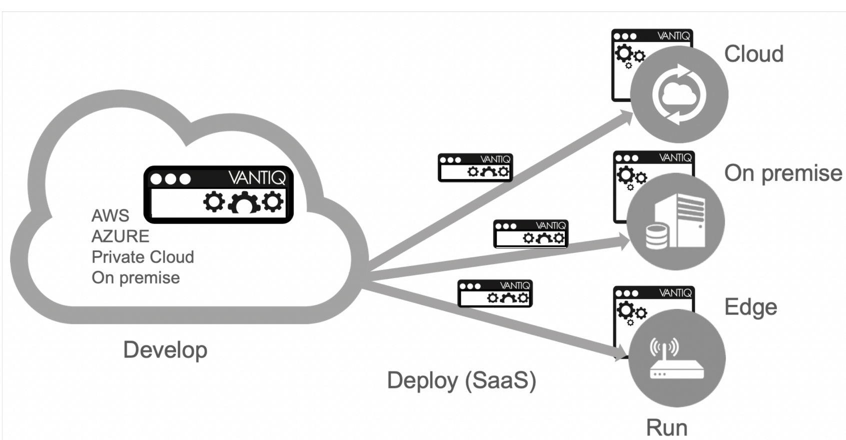
FEDERATED DISTRIBUTION – MULTIPLE TOPOLOGIES



Uniform Application Model Across the Distributed Environment

Finally, it is worth noting that every node in a VANTIQ network implements the complete set of VANTIQ services. Utilizing the VANTIQ deployment managers system functionality can be dynamically re-distributed throughout the VANTIQ network to optimize the tradeoffs of data communication, latency, processing efficiencies and fault tolerance.

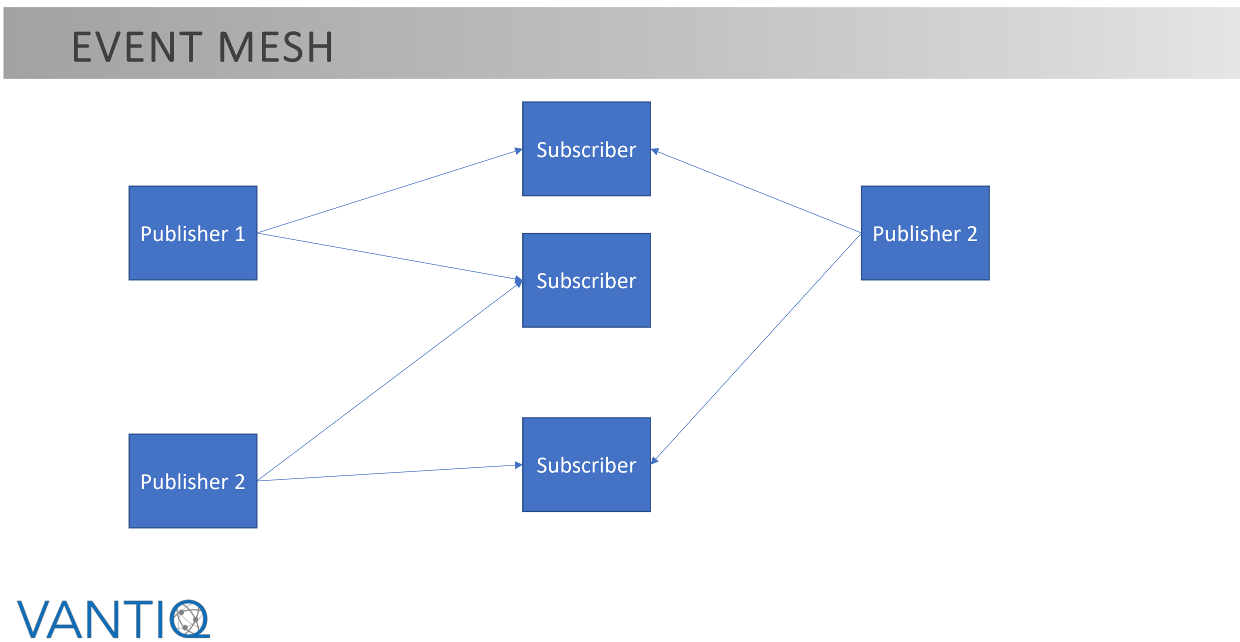
DISTRIBUTION ACROSS MULTIPLE INSTALLATIONS



VANTIQ Event Mesh

VANTIQ employs a unique approach to event-driven integration that serves to improve scalability, availability and to reduce operational costs for managing the event-driven integration topology by taking advantage of VANTIQ's federated, distributed architecture.

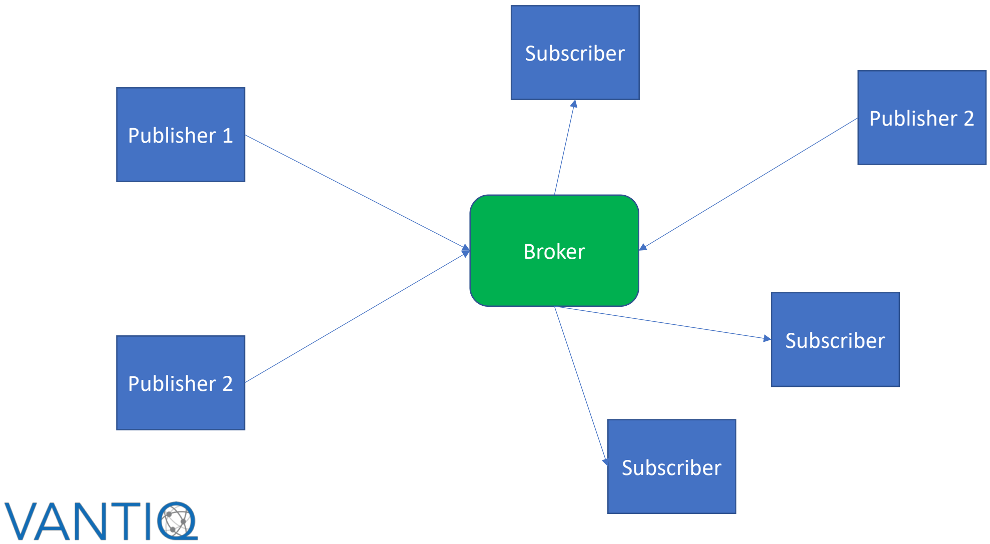
Most Advanced Event Brokers typically use a standard pub/sub model. A diagram of a prototypical event mesh from the logical perspective:



Both publishers and subscribers are applications. The goal is to be able to use events published by one application in other applications. This allows the applications to work together as systems to solve larger and more strategic business problems while still maintaining a high degree of agility due to the asynchronous, loosely-coupled nature of the pub/sub model and the simplicity of the event payloads. Publishers and subscribers can evolve completely independently as long as the event payloads remain upward compatible or, in the case of Pronto, as long as the event broker can transform the events to maintain upward compatibility.

However, if you look at the physical implementation of a pub/sub network using an advanced event broker, the implementation could be based on a centralized broker resulting in a physical implementation that is, in reality, a star topology:

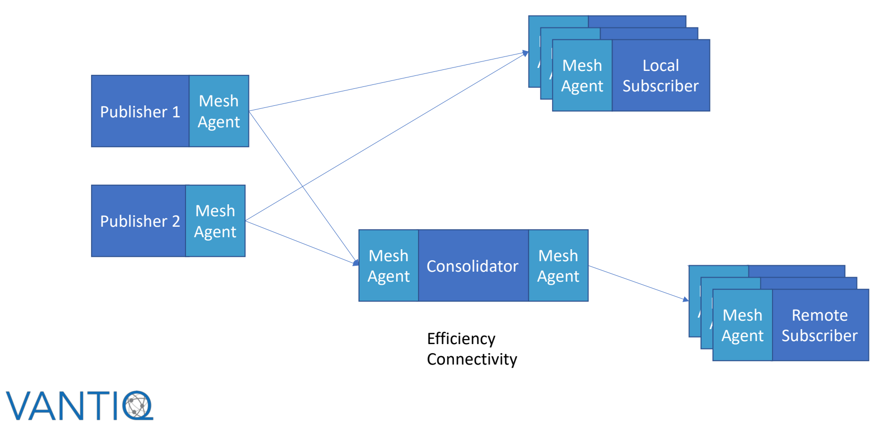
EVENT MESH WITH CENTRALIZED BROKER



This topology may be the source of scalability and reliability problems. If nothing else, tuning such a broker for high scalability is a difficult task.

An alternative to a physical implementation using the centralized broker approach is the VANTIQ brokerless event mesh. In this topology events are delivered directly from publishers to subscribers rather than requiring the events to pass through the centralized broker. This simplifies the problems of scaling and availability since there is no bottleneck or single point of failure for the event mesh. This alternative is depicted graphically as:

VANTIQ BROKERLESS EVENT MESH



Summary

Connected Industry exploits all forms of integration to achieve its business goals:

- Event-driven – coordinate activities throughout the business in real-time
- Service-driven – construct each application in an agile fashion using micro-services. The micro-services implicitly support sharing data among applications
- Data-driven– share data among applications using asynchronous and batch integration approaches
- Collaboration – efficiently integrating people with the intelligent systems

with the emphasis on next generation real-time event-driven integration and collaboration.

In addition, simple connectivity is not enough to solve the integration problem in an agile and efficient manner. The integration platform must be capable of real-time analysis using time series, statistical and machine learning methods and be able to respond to the situations identified by the analysis by effectively collaborating with the users and operators of the systems.

Finally, all forms of integration must be construction with geographic distribution in mind as well as the need for the geographically distributed systems, controlled by many different organizations, to be efficiently federated to provide the needed levels of integration as well as the local control and security required by independent, perhaps competing, organizations.

VANTIQ provides a state of the art, real-time, event-driven integration platform supporting the integration of devices, machines, applications, AI and people into an effective set of federated systems that enable next generation, real-time Connected Industry.

Appendix A: Integration Architectures Overview

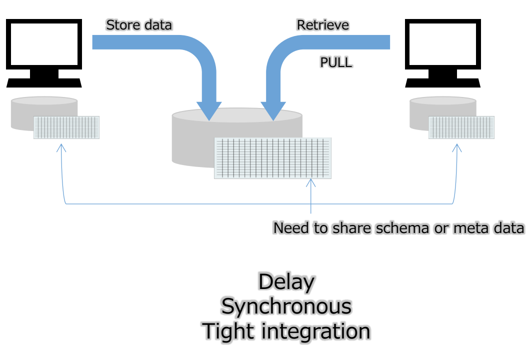
It is expected that all forms of integration will have a role in Connected Industry but that real-time, event-based integration will serve as the foundation of the initiative. A brief overview of the most common forms of technical integration is presented so that the reader will be familiar with the supporting forms of integration and their role in Connected Industry.

Data-driven Integration

Data-driven integration is the traditional form of integration that has been used to integrate the data produced and consumed by multiple systems. Data integration implements the following strategies:

- File transfer – periodically, perhaps once a day or once a week, a file is produced that contains the desired data. This file is then transferred to another location, possibly transforming the data during the transfer. Once transferred, the data is available to the requesting application. A variant of file transfer is the database replication model that transfers changes from one database to another.
- Messaging – the requesting application sends a request to the owner of the data. The owner finds the requested data and sends a message back to the requestor. The requestor receives the data and incorporates it into their application.
- Request-response – a synchronous request for data from another system. This may take the form of a request to the system or a read on the database effectively causing the database to be shared by the two systems. This is the traditional way databases are used to provide a shared store for multiple applications.

In most cases, the integration is not real-time involving some delay before the data is obtained by the requestor. Even the synchronous cases incur indirect delays because the write of the data to the database had to precede the request by some, many times significant, amount of time. This makes traditional data integration suboptimal for real-time connected industry applications.



Event-driven Integration

Event-driven integration is a new integration architecture enabled by Advanced Event Brokers and the event mesh. Using this architecture, the applications that produce events publish them directly to the advanced event broker. Applications that wish to be notified of the business events subscribe and the advanced event broker delivers the events to the subscribers in real-time. There is no delay while data is written to a database and then, subsequently, retrieved by the interested application.

Event-driven integration has a number of advantages over traditional data integration:

- It is naturally real-time enabling real-time business and, in this case, real-time Connected Industry.
- Event-driven integration is naturally loosely-coupled as all integration is via asynchronous publish/subscribe architectures. This completely decouples the publishers from the subscribers. This loose-coupling drastically increases agility in Connected Industry applications and systems.
- The payload associated with an event is usually quite simple making its use in subscribing applications relatively simple.
- In addition to simple data models, the schema for the published event payload is usually much more stable than typical data integration approaches using databases, file transfers and messaging system where the data integration schema is commonly based on the publishing application's database schema. This means that the events are less likely to suffer incompatibilities when the publishing application and its database schema are changed. The fact the event payload is stable also makes the overall system supporting connected industry much more agile as changes can be made in publishers while rarely impacting subscribers.

Real-time
Asynchronous
De-coupled between applications

Service-driven Integration

A final form of integration is a more modern synchronous data integration model but using micro-services as the suppliers of the requested data. Service integration is currently a popular topic due to the proliferation of micro-service-based architecture initiatives.

Micro-services are very useful for synchronous access to business logic but far less useful for integration in real-time connected industry:

- Calls are synchronous reducing agility as the two systems must be bound at execution time.
- Requestors looking for real-time data generally have to poll the micro-services requesting updates to the data. This results in a lag between the time the data is produced as an event and

the time the requestor can be notified of its availability. This results in either wasted resources due to low duty cycle polling or significant lag time with more efficient high duty cycle polling.