Issue 28 - December 2016

SkyFoundry Insider

The World of Data Analytics for the Internet of Things, Buildings and Equipment Systems Just Changed...Forever

A Special Issue - Introducing...

SkySpark EVERYWHERE

Introducing the SkySpark® Everywhere™ Distributed Informatics Architecture

Now You Can Apply SkySpark from the Edge to the Cloud

The Internet of Things is a Distributed Computing Challenge

The IoT has gained widespread focus and a growing volume of hype. Those of us in the automation industry can rightfully say that we have been connecting devices to networks and the Internet for decades. So what is different now?

"The reality is that it is not possible, cost effective or desirable to transmit every piece of data from every IoT device to the cloud in order to gain value from that data.

An IoT technology platform needs to recognize and embrace the highly distributed and innately non-hierarchical nature of the IoT and support that with a corresponding software architecture."

Smart Systems Design | Harbor Research

There are a number of important factors driving change:

- Lower costs – new price points make new IoT products and applications possible
- Entrants from outside of the automation industry are challenging industry norms, price points and methods
- Recognition by society at large that connected devices can provide desirable benefits
- Intersection with societal factors – sustainability, personal devices, ubiquity of networks
- **The importance of DATA**

The value of data from connected devices may be the most important factor driving change and acceptance of the IoT. Only by connecting to devices can we realize the value contained in their data – to inform operators, improve processes, insure safe environments and optimally manage resources. Without easy access to data the promise of the IoT cannot be realized.

That's easy to say but the reality is that connecting to data from diverse devices, managing it and deriving value from it is a technical challenge that spans a range of technologies and the often isolated communities that have grown up around them.

One of the most fundamental characteristics of the IoT is that it is a distributed computing challenge. Similar to how Building Automation Systems started in the 1980's with centralized computer systems and then moved to distributed-control architectures, the IoT is distributed computing challenge that requires a fundamentally different technology architecture - **SkySpark Everywhere™ takes you there.**

Con't on pg 3

Now Apply SkySpark from the Edge to the Cloud

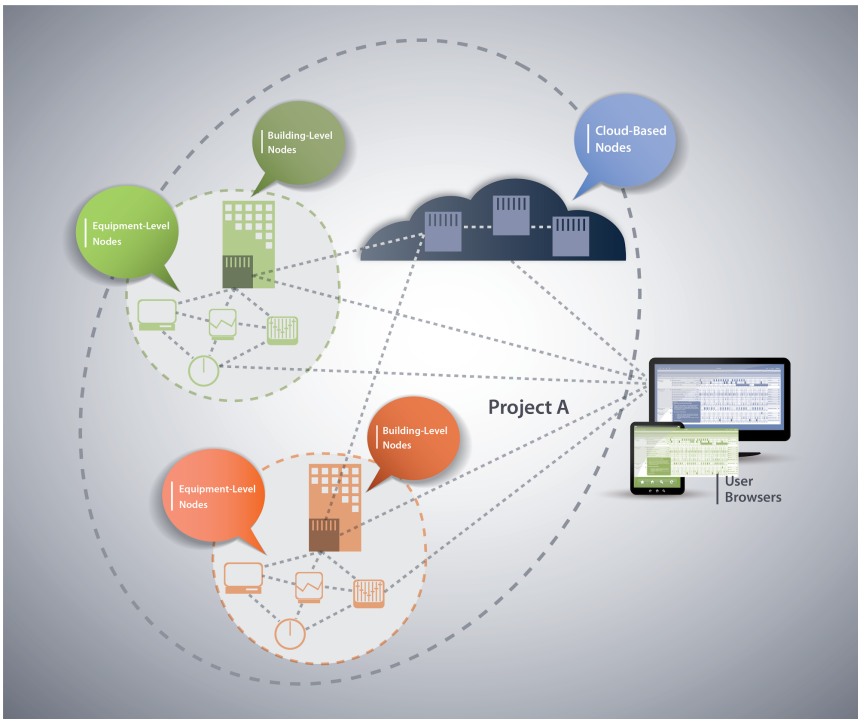
Introducing the SkySpark Everywhere Distributed Architecture & Pricing Model

Con't from pg 2

We hear lots of talk about the "cloud" as it relates to the IoT. In many cases it seems like the "cloud" is being presented as the solution to all things IoT. The reality is that it is not possible, cost effective or desirable to transmit every piece of data from our devices to the cloud in order to gain value from them. A true IoT architecture needs to recognize and embrace the highly distributed nature of the IoT and provide a software architecture that matches that reality.

Software functions need to reside at every level of the architecture from the "edge" (on an equipment system or in an electrical closet) - to the building level - where data from multiple smaller nodes can be aggregated and analyzed, to the portfolio level where data and analysis occur at the server or cloud level.

SkySpark Everywhere software has been designed to run in everything from a small low cost device the size of a deck of cards, to server clusters hosted in the cloud. Throughout the architecture SkySpark provides the exact same software architecture, data model, functions and features. There is only one platform to learn and one set of tools across to work with. The result of this unified end-to-end software architecture is faster development, an easier learning curve, more reliable applications and a streamlined near seamless user interface across the entire data architecture.



SkySpark Everywhere Physical Architecture View

All New Pricing Too. SkySpark Everywhere's technology is one key part of this transformative solution, the other is the new pricing model that lets you apply SkySpark in capacities **as small as 10 points** with a **price that starts at $60 USD** ! And make no mistake, this includes **ALL of the features of SkySpark,** not some toy feature set or simple data collection agent.

Data Democracy – Equal Treatment and Services for All Types of Data, Protocols and Methods of Acquisition

Data of All Types and Formats

The IoT is characterized by very different types of data that come from different sources and are represented in many different formats. While we might wish for a singular data standard, we are far better off to create technology that works with the realities of diverse, multi-structured data. That starts with the ability to communicate via numerous open, standard communications protocols, and the tools to create additional communication connectors as needed.

Data Acquisition – Communications Connectors. In some cases, data is acquired via live links streaming real-time or near real-time data. In other applications data is acquired and processed in batches that contain historical sensor data – think CSV files, Excel sheets or queries to external databases. Whether via a live link to an automation system or smart meter, connection to an SQL database, import of historical data from Excel files, or a web service feed from a utility – SkySpark provides the services to consume, manage and analyze your data. SkySpark currently includes the following communication connectors:

- BACnet IP
- SNMP
- OPC UA
- Modbus TCP
- oBix
- MQTT
- Haystack
- Sedona
- SQL via JDBC
- CSV data import
- REST API

Integration With Third-party Applications

No one application does it all. Data connectivity means more than just supporting device protocols. IoT applications by nature involve an ecosystem of people, products and applications. SkySpark has been designed to enable open integration with external software applications via a fully documented and supported RESTful API.

Once we have data we need to effectively store and manage it. SkySpark's Folio database technology, developed specifically for the unique needs of IoT data, combines a number of advanced techniques to be able to accept, store, normalize and perform automated analytics on diverse data sets. Folio combines the techniques of document orientated, graph oriented and time series database design into a unique high performance database ideally suited to performing analytics and data transformations on diverse sets of sensor and machine data.

And lets talk about efficiency – Folio stores a time-stamped history sample in an average of **12 bits** (yes that's bits not bytes). Compare that to conventional database technologies that use 30, 40, even 100 bytes to store the same data sample. If you are paying for servers that difference means serious $$$ savings.

Where Functions and Services Execute Matters – *Processing Data Analytics at the Edge is Essential*

Location of Analytics Processing

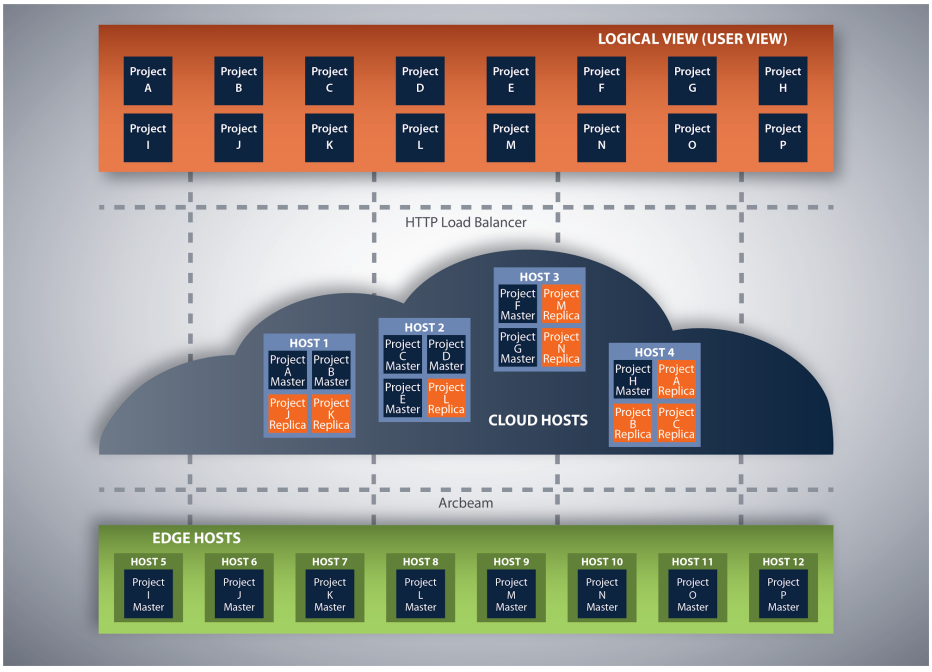
Another key consideration in implementing IoT solutions is the location of data logging and analytics processing. SkySpark is unique in that it brings the analytics processing to the data. This enables the platform to deliver high-speed continuous analysis of data – including streaming data. With SkySpark the analytics engine runs "in the data" as opposed to being a separate process dependent on delivery of data sets pulled from a database. And with SkySpark Everywhere that means that the full SkySpark software stack runs in small smart devices operating at the "edge".

What You Can Process Matters Too!

Pre-built analytic functions can be very helpful – we know - SkySpark includes hundreds of them. But the diverse world of devices and systems requires fully programmable analytics. SkySpark's Axon analytics engine enables domain experts from any field to implement rules and algorithms that fit the needs of their unique applications. SkySpark lets you leverage your expertise to add value. SkySpark provides you with the best of both worlds by combining full programmability with a library of over 500 built in analytic functions to streamline the process of implementing analytic rules and data transformation functions.

How You Interact With Your Data – Seamless Views Across Highly Distributed Systems with Hundreds to Thousands of Nodes

So now we have this highly distributed systems consisting of hundreds of connected nodes. The fact that they are distributed and not all on one server is a complexity to be hidden from the average user. Users simply want to see their data and analytic results. SkySpark Everywhere's unique technology allows the user to simply navigate their system without having to consider 'where" data or analytics processing is located. Navigation uses a unique alias-based approach all built on tagging (Haystack compatible of course!).



SkySpark Everywhere Logical Architecture

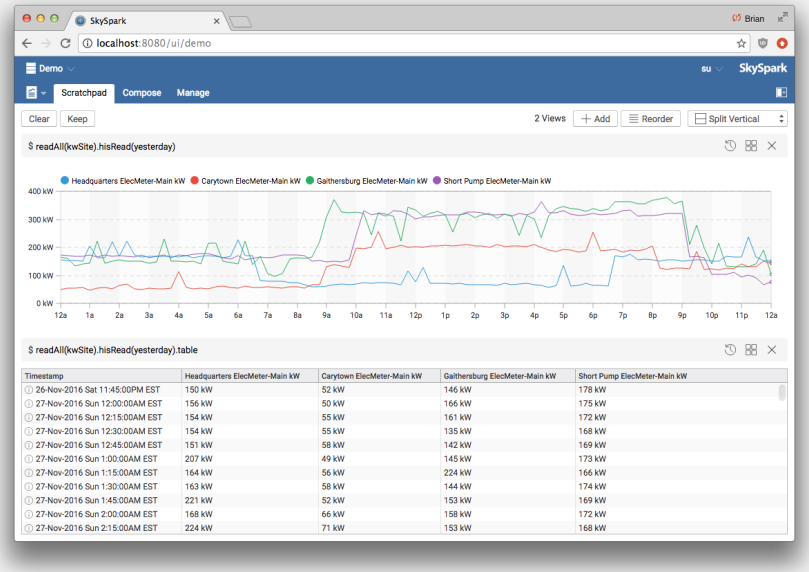
The User Experience – Bringing it All Together To Deliver Results & Value

The UX – Making Data Analytics Meaningful to Operators

Analytics results are only valuable if they drive a process to address issues to improve efficiency and reduce cost. If the process required to create those visualization results in significant costs and extra effort the value of the analytics is diminished and the potential savings will not be realized. SkyFoundry addresses this challenge with a number of unique User eXperience solutions.

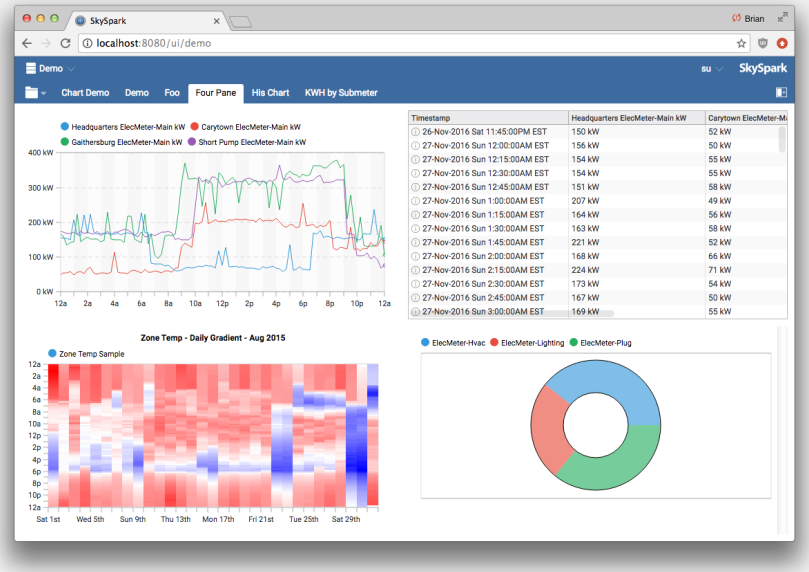
Automatic Generation of Information Visualizations

Finding patterns and transforming data may be the technological core of an analytics platform, but to make results useful we need to present those findings in visualizations that make it easy for operators to understand and act on them. SkyFoundry automatically generates rich, intuitive visualizations to show operators analytic findings, trends, correlations and relationships without the conventional approach of requiring assembly of graphical displays.

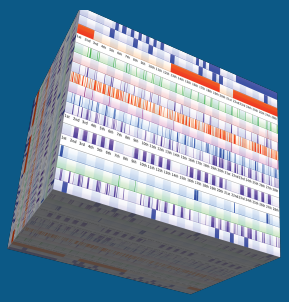


"Compose-able" Applications

Different users and applications have different needs and often want to see their data and analytics results in different ways. SkySpark combines the power of automatically generated visualizations with the ability to combine our standard tools into custom views with a feature we call compose-ability. This is not simply assembling information and charts into fixed reports, but truly combining visualizations into unique applications. SkySpark Everywhere pioneers this dramatic new advance in user experience users to create customized applications that allow user input and selection of views and data sets. We think its a revolutionary new take on web-based UI's for smart device applications.



"Automated Informatics" versus Analytics... What's in a Name?



We describe SkySpark as Automated informatics™ because it addresses the end-to-end process of generating value from data. First, it includes the communications connectors to enable you to gather data from diverse equipment systems. SkySpark supports a variety of data acquisition connectors: Bacnet IP, Modbus TCP, Obix, Haystack, SNMP, Sedona, OPC, MQTT, SQL, CSV import (manual batch or automated), and a REST API.

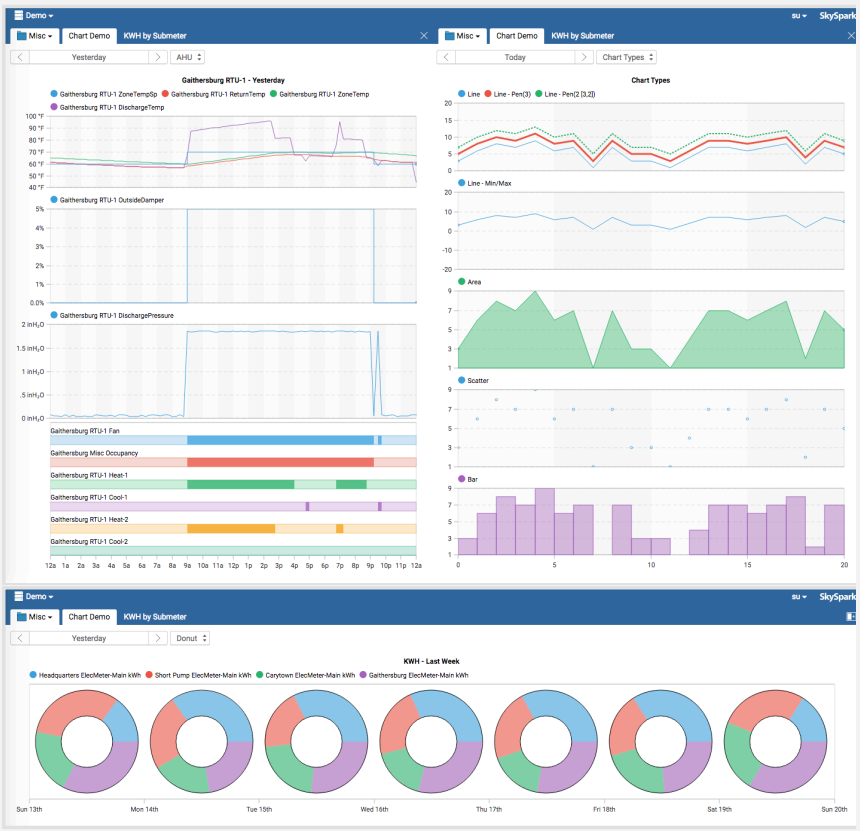
Next it includes the database technology to enable diverse, multi-structured data to be normalized and stored efficiently to make it available to higher-level applications - like the SkySpark analytics engine and visualization Apps.

A Definition

Automated Informatics: Informatics is the science of information and computer information systems. It involves the practice of information processing, and includes the interaction between humans and information.

Next it includes the high-speed analytics engine that automatically processes analytics rules against the data to identify patterns, correlations, faults, deviations and anomalies – what we call "sparks".

Then SkySpark takes the next, and possibly most important step – it automatically generates visualizations on the detected issues. SkySpark's clear, understandable, informative visualizations show operators exactly how their systems are performing and the issues that matter. Beyond finding the issues in the data, it's this step to automatically visualize analytic results that drives the workflow process to assess and correct issues.



By automatically generating rich informative displays, SkySpark makes analytics actionable and cost effective.

This unified end-to-end solution is why we call SkySpark Automated Informatics™.

So what's in a name? In this case the solution to effective, actionable analytics **making data-driven facility management a reality**.

Security is Fundamental to SkySpark® Everywhere™

Its Built Into the Architecture from the Edge to the Cloud

SkySpark is built with a **security first** philosophy. It starts with the fundamental notion that complexity breeds security holes. Security must be strong, but simple for users to implement in order to achieve a secure system. SkySpark's simple, easy to understand security architecture provides for a secure environment and utilizes multiple redundant layers of security. We incorporate the latest security practices and technology, and continually evaluate security enhancements during every release.

Network Security. Securing SkySpark's network interfaces is a critical aspect to ensure system security. SkySpark supports two different network interfaces: a **websocket** based peer-to-peer protocol named Arcbeam for communication between SkySpark nodes in a distributed system, and an HTTP interface typically utilized by users to interact with the system.

SkySpark Node-to-Node Security in Distributed Systems - Arcbeam. The Arcbeam protocol is SkySpark's peer-to-peer technology, which underpins the advanced SkySpark Everywhere distributed architecture. It is designed to work in tandem with existing security best practices such as firewalls, VPNs, and TLS encryption. Arcbeam is layered above "websockets" to establish a peer-to-peer communication link using a HTTP handshake. Once a communication link is established the connection is fully peer-to-peer, which means that **either end point can initiate the connection** without loss of any features. This makes it easy to establish distributed architectures where one endpoint is safely hidden behind a firewall. It eliminates all of the complexity of traditional IoT architectures and decisions about push-vs-pull communications links and concerns over requests for data coming from outside the network. And because it sits on top of tried and true IP infrastructure of TCP, HTTP, and Websockets it works cleanly over standard security technologies such as VPNs.

WebSocket is a computer communications protocol, providing full-duplex communication channels over a single TCP connection. The WebSocket protocol was standardized by the IETF as RFC 6455 in 2011, and the WebSocket API in Web IDL is being standardized by the W3C.

Unlike HTTP, WebSocket provides full-duplex (bi-directional) communication. Additionally, WebSocket enables streams of messages on top of TCP. TCP alone deals with streams of bytes with no inherent concept of a message. Source: Wikipedia: https://en.wikipedia.org/wiki/WebSocket)

Cryptographic Keys for Secure Authentication.
SkySpark uses cryptographic key pairs to bi-directionally authenticate each Arcbeam connection between nodes. Each SkySpark instance (node) is secured with a 2048-bit RSA public/private key pair. In order for two SkySpark nodes to establish an Arcbeam communication link, each node must be securely configured with the remote node's public key. This ensures a secure network connection is made only once each side verifies trust in the other endpoint.

Con't on pg 9

Security is Fundamental to SkySpark Everywhere™ – From the Edge to the Cloud

Con't from page 8

Arcbeam security can be further enhanced with the use of a TPM chip. TPM, trusted platform modules, provide highly secure management of cryptograph keys using a hardware chip. TPMs guarantee that access to the sensitive private keys is only available to the hardware itself, making it extremely difficult to steal private keys.

The ability for SkySpark to utilize TPM based keys further increases security and will streamline acceptance by IT security departments. Already, organizations like the DoD are specifying that new hardware products implement TPMs.

User Access. User authentication often still requires traditional username/password combinations. SkySpark utilizes a set of industry leading security technologies to manage and authenticate user passwords. Passwords themselves are stored using a technology called "PBKDF2 hashing". PBKDF2 leverages a "key stretching" algorithm, which requires significant CPU cycles to compute a hash of a password. This raises the difficulty of guessing the original passwords even if the system is compromised.

Authentication of username and passwords uses another cyber security standard - SCRAM (Salted Challenge Response Authentication Mechanism). SCRAM leverages the PBDKF2 hash to ensure that neither the original password, nor the PBDKF2 hash are sent over the network. It also allows both client and server to verify that they agree on the password.

Arcbeam - Highly Efficient Message Protocol Minimizes Bandwidth

Arcbeam messages are encoded using Brio, our new data compression technology. This is the same technology that makes it possible for a time stamped value to be stored in 12 bits in the SkySpark's Folio database. In Arcbeam it compresses the entire message payload. The result – less latency and less bandwidth utilization on networks. Another feature that IT departments look for.

From network communication to user management, cyber security is a core part of the foundation of SkySpark Everywhere.

Taking Security Keys to the Next level – Integration with TPM – the Trusted Platform Module

SkySpark's Arcbeam implementation enables keys to be based on the highly secure TPM technology.

Trusted Platform Module (TPM) is an international standard for a secure cryptoprocessor, which is a dedicated microcontroller designed to secure hardware by integrating cryptographic keys into devices. TPM's technical specification was written by a computer industry consortium called Trusted Computing Group (TCG).

Many new IoT devices now include a TPM. Since each TPM chip has a unique and secret RSA key burned in as it is produced, software (like SkySpark) can use the Trusted Platform Module to generate unique certificates to authenticate hardware devices and communications across clusters of nodes.

Source: Wikipedia: https://en.wikipedia.org/wiki/Trusted_Platform_Module

Its On !!!!

Haystack Connect 2017 has been announced for May 2017 !!!
The conference and exhibition will take place at the fantastic Saddlebrook resort near Tampa FL. Attendee, Exhibitor and Sponsor registration and the Call for Speakers is open now! (And expect the IBB to be back too)

You can find full details here: https://haystackconnect.org/

Take advantage of Early Bird pricing through Jan 31, 2017!!!