



*"Find what matters"™*

## SkySpark® Everywhere™ – Introducing the Distributed Informatics Architecture

*Now Apply Analytics from the Edge to  
the Cloud in a Unified, Distributed System*



**CONTENTS**

Introduction .....2

Introducing the SkySpark® Everywhere™  
Distributed Informatics Architecture.....3

Benefits for Real World IoT Applications .....3

Options for Deployment at the “Edge” .....4

Where Functions and Services Execute Matters  
for IoT Applications .....5

Understanding Distributed Architectures with  
SkySpark Everywhere .....6

Interacting With Your Data - Seamless Navigation Across  
Distributed Systems .....8

The User Experience – Bringing it All Together To  
Deliver Results & Value.....9

Security - Built Into the Architecture from the Edge  
to the Cloud .....11

SkySpark® – Analytics for a World of Smart Devices .....13



# The Internet of Things is a Distributed Computing Challenge



## Introduction

The Internet of Things (IoT) has gained widespread focus with increasing adoption driven by factors including:

- 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 is perhaps the most important factor driving industry 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.

**“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.”**



That’s easy to say, but the reality is that connecting to data from diverse, widely dispersed devices, managing that data and deriving value from it is a technical challenge that spans a range of technologies.

One of the most fundamental characteristics of the IoT is that it is a *distributed computing challenge*. Similar to how computer-based automation systems started in the 1980’s with centralized computers and then moved to distributed-control architectures, the IoT is distributed by its very nature and requires software technology that supports that reality. *SkySpark® Everywhere™ provides that technology.*

## Introducing the SkySpark® Everywhere™ Distributed Informatics Architecture

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 the data. A true IoT architecture needs to recognize and embrace the highly distributed nature of the IoT and enable data processing at the edge as well as the cloud. An IoT software platform must provide a software architecture that matches that reality.

Software functions need to be able to reside at every level of the architecture from the “edge” – on an equipment system, in an electrical closet or in a vehicle, to the building level – where data from multiple smaller nodes can be aggregated and analyzed, to the cloud where data analysis occurs at a portfolio level. SkySpark Everywhere provides this capability.

SkySpark Everywhere allows systems of 2 to hundreds of SkySpark nodes to be connected in unified, near-seamless systems. Data collection, storage, management, analytics processing and visualization functions are performed by the individual nodes at whatever level they are deployed, while user applications that present data and analytics results to users work as if the system were one, single unified server.

### Benefits for Real World IoT Applications

By implementing a true distributed architecture, SkySpark Everywhere provides numerous benefits, which directly address needs of IoT applications including:

**Local data collection ensures reliability.** Network issues, can interrupt data collection in cloud-dependent systems, resulting in loss of critical data. Data collection services operating at the edge insure that data is collected by the local node where the risk of communication issues is far lower. The system still provides for synchronization of data (via replication) to enable access at the cloud level without the need for real time communication to edge devices.

**Lower data transfer costs.** In addition to reliability, SkySpark Everywhere’s distributed architecture provides optimized synchronization of data with the cloud, which

#### SkySpark Everywhere – A Pricing Model to Support Use at the Edge

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. This includes ALL of the features of SkySpark, not a reduced feature set or simple data collection agent.

can dramatically reduce data transfer costs in systems that rely on cellular networks for continuous transmission of data to the cloud for processing. Typical user requests for data and visualizations are directed to the most recent “replica” instead of initiating requests over the network. The reduction in network usage can be dramatic.

**Simpler implementation.** With only one platform to learn and one set of software tools to work with, the result is faster development, an easier learning curve, and more reliable applications.

**A Simpler user experience.** The fact that data and processing is distributed across multiple nodes is a complexity to be hidden from the average user. SkySpark Everywhere allows them to navigate their world as if it were a singular system. They choose to look at projects, sites, buildings, or vehicles – navigating by whatever logical entities make sense for their application. SkySpark Everywhere provides a streamlined, near seamless user interface across distributed systems.

## Many Options for Deployment at the “Edge”

In order to be effectively applied at the edge costs have to be competitive. 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.

With the advent of new IoT gateway hardware options, systems integrators and system designers have greater choice in the hardware they use, reducing risk of vendor lock-in and insuring competitive pricing. With minimum processor specifications similar to those shown, SkySpark Everywhere can be deployed on a wide range of standard, easily available, low cost edge devices.

### SkySpark Everywhere Minimum Edge Hardware Specification

1 GHz processor  
512 MB RAM for up to 100  
points

Our test platform:  
<http://beagleboard.org/black>

Software: Linux OS with a Java  
Virtual Machine

## Where Functions and Services Execute Matters for IoT Applications

### *SkySpark Everywhere Brings Data Collection, Analytics and Visualization to the Edge*

A key consideration in implementing IoT solutions is the location of data collection, logging, storage and analytics processing. Real-time applications cannot be served with the latencies involved in transmitting data to the cloud for processing and analytics.

SkySpark Everywhere enables data collection, storage, management, analytics and visualization to be performed by the edge device. Operators can view comprehensive data and analytic presentations served directly by SkySpark operating in the edge node. They do not have to access a cloud host outside of their facility. This capability can be an important factor to address security concerns in mission critical applications, as well as real time applications.

SkySpark is also unique in that it brings the analytics processing to the data. With SkySpark the analytics engine runs where the data source is as opposed to being a separate process dependent on delivery of data sets pulled from a database. This enables the platform to deliver high-speed, continuous analysis of data – including streaming data.

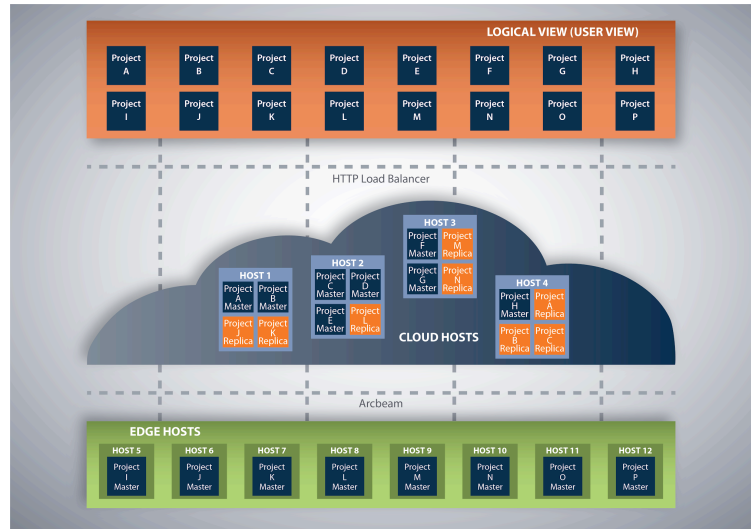
## What You Can Process Matters Too – With SkySpark You get the Power of Fully Programmable Analytics

Pre-built analytic functions can be very helpful but are not sufficient for real world applications. It's essential that domain experts can write rules and algorithms that fit the unique needs of their applications. The diverse world of devices and systems requires fully programmable analytics. SkySpark's Axon analytics engine enables experts from any field to implement rules and algorithms that fit the needs of their unique applications.

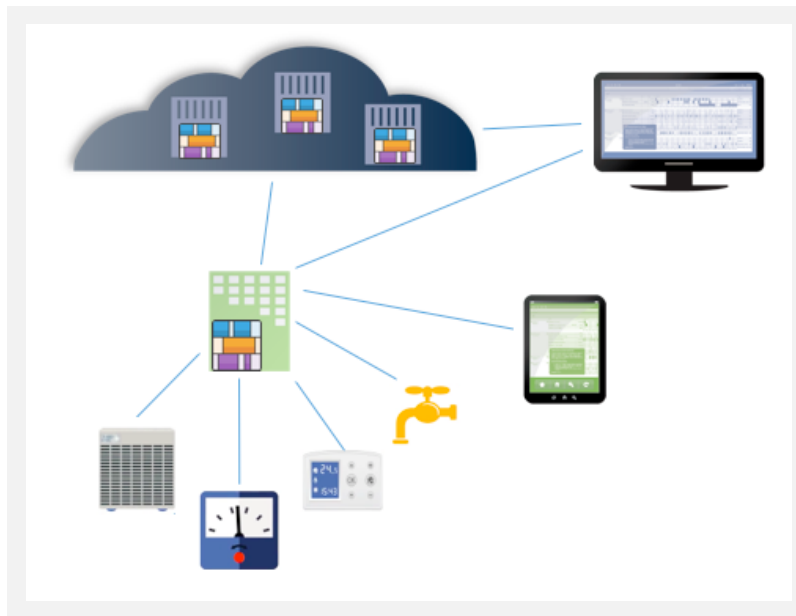
SkySpark combines full programmability with a library of over 500 built in data analytic functions to streamline the process of implementing analytic rules and data transformation functions – *an approach that provides the best of both worlds*. This white paper provides more detail on this subject: <https://skyfoundry.com/file/48/SkySpark---Combining-Full-Programmability-and-Built-in-Analytic-Functions.pdf>

## Understanding Distributed Architectures with SkySpark Everywhere

SkySpark Everywhere provides tremendous flexibility in the design of system architectures to achieve the optimal solution for your application. Below we review some sample architectures.



### SkySpark Everywhere deployed on-premise at the building level



In this architecture, a node running SkySpark is installed in the facility and connected to local data sources via a variety of standard data acquisition connectors including: Bacnet IP, Modbus TCP, Obix, Haystack, SNMP, Sedona, OPC UA, and the REST API – all included with SkySpark.

In this approach SkySpark is acting as an in-building, facility-level data aggregator, yet provides features far beyond simple data collection and

gateway functions by including all of the capabilities of SkySpark. Data and analytic results are accessible **at the facility** without requiring users to access a cloud-based server to view their data. Users outside of the building network can still access the data at the cloud level through SkySpark's replication features, which synch data on a configurable basis to insure compatibility with network limitations.

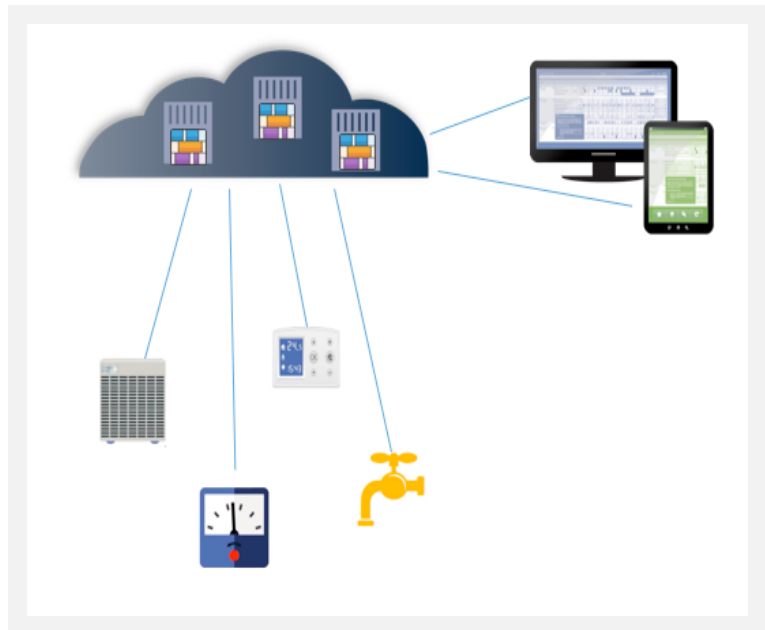
### **SkySpark Everywhere deployed on-premise in edge devices on equipment systems**

In this architecture SkySpark is installed on small “edge computing” platforms (IoT gateways and OEM controller products) that directly manage equipment systems, or are connected to other smart devices or sensors. Every node can provide local user access to all of the capabilities of SkySpark.



### **Conventional cloud-centric architecture with data collected from external devices by SkySpark nodes hosted in the cloud**

This architecture is a good fit in applications where the end devices have local data collection and logging capability. Communication to SkySpark instance(s) operating at the cloud-level typically employs “synching” of trend logs stored in the end device. Here, clustering allows large-scale systems to be distributed over multiple servers for load balancing and performance optimization. “Live polling” of data from end devices is also supported in this architecture, but the application designer needs to consider the effects of network latency, bandwidth, communications faults and the ability of end systems to support the necessary volume of polling requests.





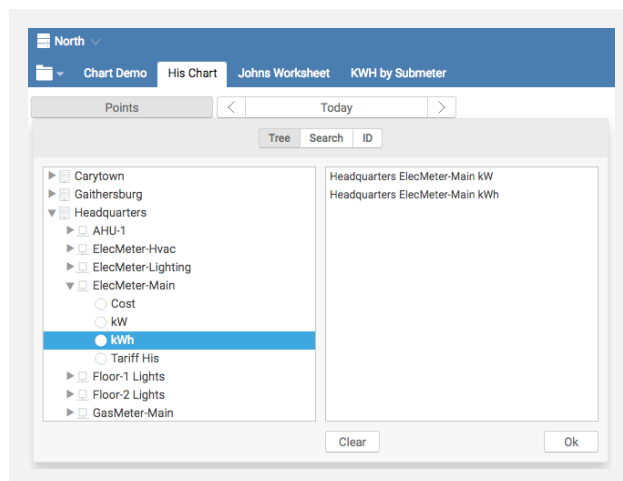
## Interacting With Your Data - Seamless Navigation Across Distributed Systems

So once we have this highly distributed system consisting of many connected nodes, ease of user interaction is the key requirement. The fact that the data and analytic computation is distributed and not hosted 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 navigate their system without having to consider 'where' data or analytics processing is located. They choose to navigate by whatever logical entities make sense for their application – projects, sites, buildings, or even vehicles in mobile asset applications.

SkySpark Everywhere provides a streamlined, near seamless user interface across the entire distributed data architecture via a standard browser. Views are presented using HTML5 technology – there are no apps, clients or plug-ins to load. Navigation uses a unique alias-based approach built on the concept of tagging. (SkyFoundry follows the Project-haystack.org tagging methodology standard).



Navigation can be tailored to any logical model of "entities" such as projects, regions, campuses, sites, floors, and equipment hierarchies. How the data is spread across a cluster of nodes is completely hidden from the operator - they simply navigate data in a way that makes sense for their application. Entities could be as small as a few sensors on an embedded node, or as large as all of the environmental systems in a million sq foot office building.



User-defined entities appear in navigation trees and selection boxes. This approach enables designers to "model" system in a way that fits the user's view of the their application.

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

## The User eXperience – Making Data Analytics Meaningful to Operators

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

## Automatic Generation of Information Visualizations

Transforming data and finding patterns – faults, anomalies, deviations – 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.

SkySpark automatically generates rich, intuitive visualizations to show operators analytic findings, trends, correlations and relationships without the conventional approach of requiring assembly of graphical displays.

And with SkySpark Everywhere these rich visualizations are delivered directly from even the smallest node.

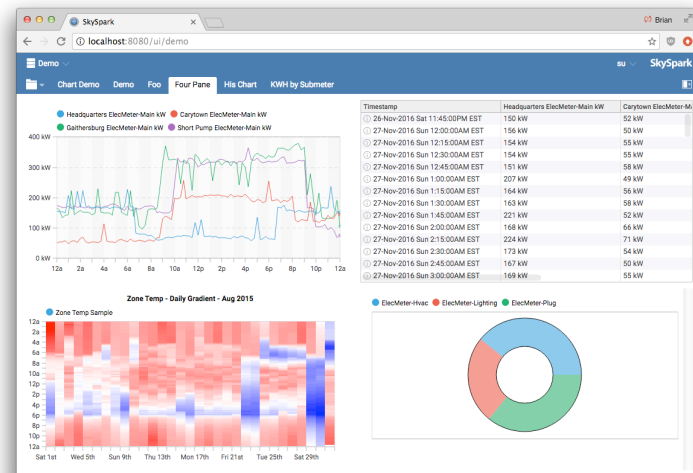
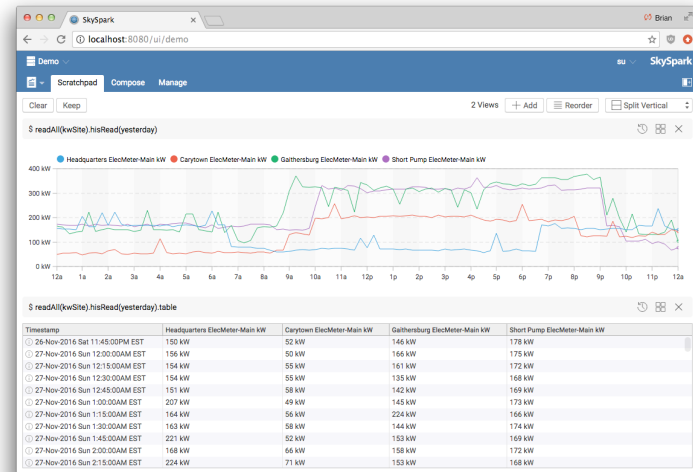


## “Compose-able” Applications

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 allowing users to create customized applications that allow user input and selection of parameters and data sets – all without requiring software development skills.

We think it’s a revolutionary new take on web-based UI’s for smart device applications.



## Security is Fundamental to SkySpark® Everywhere™

### *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 node is secured with a 2048-bit RSA public/private key pair. In order for two SkySpark nodes to establish an Arcbeam 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.

Arcbeam security can be further enhanced with the use of a TPM chip. TPMs, provide highly secure management of cryptographic 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 to utilize TPMs further increases security and streamlines acceptance by IT security departments. Already, organizations like the DoD are specifying that new hardware products use 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 PBKDF2 hash to ensure that neither the original password, nor the PBKDF2 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 an average of 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 core to 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 an 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](https://en.wikipedia.org/wiki/Trusted_Platform_Module)

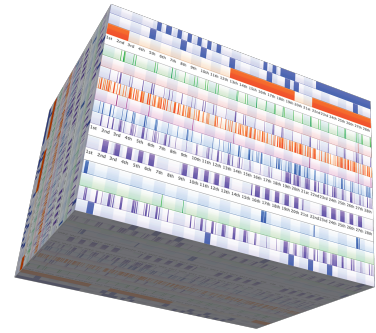
## SkySpark® – Analytics for a World of Smart Devices

The past decade has seen dramatic advances in automation systems and smart devices. From IP connected systems using a variety of standard protocols, to support for web services, and the emergence of the IoT, it is now possible to access the data produced by the wide range of devices found in today's smart devices and equipment systems.

Access to this data opens up new opportunities for the creation of value-added services to help businesses reduce energy consumption and operational costs and to identify opportunities to enhance operations through improved control, and replacement or repair of capital equipment.

Access to the data is just the first step in that journey, however. The new challenge is how to manage and derive value from the exploding amount of data available from smart and connected devices. SkyFoundry's SkySpark® directly addresses this challenge.

***The new frontier is to efficiently manage and analyze data to Find What Matters™.***



Contact us at:

[info@skyfoundry.com](mailto:info@skyfoundry.com)

Learn more at:

[www.skyfoundry.com](http://www.skyfoundry.com)



SkyFoundry's mission is to provide software solutions for the age of "the Internet of things". Areas of focus include:

- Building automation and facility management
- Energy management, utility data analytics
- Remote device and equipment monitoring
- Asset management

SkyFoundry products help customers derive value from their investments in smart systems. Learn more at:

[www.skyfoundry.com](http://www.skyfoundry.com)