Using Data Science to Proactively Manage the Connected C-Store

Presenters:
Ashwin Swamy, Omega ATC
Thomas Duncan, Omega ATC
Agenda

• Housekeeping
• Presenters
• About Conexxus
• Presentation
• Q & A
Housekeeping

This webinar is being recorded and will be made available in approximately 15 days.

• YouTube (youtube.com/conexxusonline)
• Website Link (conexxus.org)

Slide Deck
• Survey Link – Presentation provided at end

Participants
• Ask questions via webinar interface
• Please, no vendor specific questions
• Our webinars may be used toward PCI continuing education credits. Please contact arussell@conexxus.org for questions regarding a certificate of webinar attendance.

Email: info@conexxus.org
Presenters

Conexxus Host
Allie Russell
Conexxus
arussell@conexxus.org

Moderator
Kara Gunderson
Chair, Data Security Committee
POS Manager, CITGO Petroleum
kgunder@citgo.com

Speakers

Ashwin Swamy
Director of Cyber Resilience, Omega ATC
ashwin.swamy@omegaatc.com

Thomas Duncan
Senior Security Strategist, Omega ATC
thomas.duncan@omegaatc.com
About Conexxus

- We are an independent, non-profit, member driven technology organization
- We set standards…
  - Data exchange
  - Security
  - Mobile commerce
- We provide vision
  - Identify emerging tech/trends
- We advocate for our industry
  - Technology is policy
<table>
<thead>
<tr>
<th>Month/Date</th>
<th>Webinar Title</th>
<th>Speaker</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 24, 2019</td>
<td>Who’s Watching Your Network? What you should know about Managed Detection &amp; Response (MDR)</td>
<td>Mark Carl Tom Callahan</td>
<td>ControlScan</td>
</tr>
<tr>
<td>March 21, 2019</td>
<td>Proactive Defense in Depth</td>
<td>Brett Stewart DeWayne Mangan Mark Palmer</td>
<td>Acumera</td>
</tr>
<tr>
<td>June 27, 2019</td>
<td>Web Payment Aspirations</td>
<td>Ian Jacobs</td>
<td>W3C</td>
</tr>
<tr>
<td>July 25, 2019</td>
<td>Skimming</td>
<td>Linda Toth Paige Anderson Caleb Burke</td>
<td>Conexxus NACS CITGO</td>
</tr>
<tr>
<td>August 8, 2019</td>
<td>Application Security 101</td>
<td>Denis Sheridan</td>
<td>Synopsys</td>
</tr>
<tr>
<td>Month/Date</td>
<td>Webinar Title</td>
<td>Speaker</td>
<td>Company</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>August 29, 2019</td>
<td>Don’t Get Phished!! Train Your Employees To Avoid Ransomware</td>
<td>Geoffrey Vaughan Ed Adams</td>
<td>Security Innovation</td>
</tr>
<tr>
<td>September 26, 2019</td>
<td>Using Data Science to Proactively Manage the Connected C-Store</td>
<td>Ashwin Swamy Thomas Duncan</td>
<td>Omega ATC Omega ATC</td>
</tr>
<tr>
<td>October 24, 2019</td>
<td>Easy PCI—How to Make PCI &amp; Attestation Easier</td>
<td>Ajith Edakandi</td>
<td>Hughes</td>
</tr>
<tr>
<td>November 21, 2019</td>
<td>Applicability of PCI DSS requirements for Merchants and MNSP’s</td>
<td>Sam Pfanstiel</td>
<td>ControlScan</td>
</tr>
<tr>
<td>December 5 and 12, 2019</td>
<td>EMV</td>
<td>TBD</td>
<td>TBD</td>
</tr>
<tr>
<td>January, 2020</td>
<td>TBD</td>
<td>TBD</td>
<td>Cybera</td>
</tr>
</tbody>
</table>
NACS Show
October 1-4, 2019
Atlanta, GA

Booth # 3755

Conexxus thanks our 2019 Annual Diamond Sponsors!

mSHIFT.
Stuzo
DN
Wex
Hughes
GILBARCO
VEEDER-ROOT
Cybera
SECURITY
INNOVATION
Conexxus
Using Data Science to Proactively Manage the Connected C-Store
Using Data Science to Proactively Manage the Connected C-Store

- Difficulties with managing connected and IoT devices.
- Using industry analogs to navigate a digital transformation.
- Use cases for more effectively managing the connected petroleum retail environment.
- Tips for experimenting with predictive maintenance approaches.
The Fourth Industrial Revolution is in its earliest stages.

4IR
1. Big Data
2. Internet of Things
3. Artificial Intelligence
4. Robotics
5. Autonomous Transport
6. Augmented Reality

- Gray Taylor
  (September 2018)
The Fourth Industrial Revolution is in its earliest stages.

4IR
1. Big Data
2. Internet of Things (IoT)
3. Artificial Intelligence
4. Robotics
5. Autonomous Transport
6. Augmented Reality

- Gray Taylor
(Sepetember 2018)

The 4 Industrial Revolutions (by Christoph Roser at AllAboutLean.com)
IoT adoption is in a holding pattern.

”...concerns over integration issues—in particular, technical expertise, data portability and transition risks—have become more acute over the past two years.”

- Bain IoT Customer Survey, 2018
IoT adoption is in a holding pattern.

”...concerns over integration issues—in particular, technical expertise, data portability and transition risks—have become more acute over the past two years.”

- Bain IoT Customer Survey, 2018

Percentage of industrial customers citing these barriers as a top concern

Notes: Industrial segments include discrete manufacturing, process industries, production sites, building, infrastructure and utilities
Sources: Bain IoT customer survey, 2016 (n=533, industrial customers=182); Bain IoT customer survey, 2018 (n=627, industrial customers=329)
The prospect of more devices is not always met with enthusiasm.
The prospect of more devices is not always met with enthusiasm.
More devices out in the field means more servicing.

"I meet with the service team multiple times a week and get daily updates on the reliability of the vehicle. We -- the best service, of course, is no service. Like that's the vehicle just -- reliability and quality being so good that service is rarely required. That's what the main goal is like, eliminate the need for service."

- Elon Musk (July 26, 2019)
The predictive maintenance paradox

“Predictive maintenance is the prime example of a use case that vendors are ready to deploy, but customers are less excited about, as shown in the middle right panel. Its decline in attractiveness indicates customers face implementation barriers.”

- Bain IoT Customer Survey, 2018
The predictive maintenance paradox

We desire the ability to determine exactly when and why equipment or operations are at risk of failing in order to reduce overall maintenance costs.

IoT devices inherently increase the number of failure points in the retail environment, and therefore can increase maintenance costs.

In order to identify these signals, we require IoT devices to provide us with the data we need.
Weighing the business case vs transition costs...

<table>
<thead>
<tr>
<th>BUSINESS CASE</th>
<th>TRANSITION COSTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitigate operational issues</td>
<td>More frequent site visits</td>
</tr>
<tr>
<td>Easier Remote Management</td>
<td>Management hassle (another thing to deal with)</td>
</tr>
<tr>
<td>Better Customer Experience</td>
<td>Increase in number of operational issues and device failures</td>
</tr>
<tr>
<td>Safety improvements</td>
<td>Network issues</td>
</tr>
<tr>
<td>Reduced device failure rates</td>
<td>Security issues</td>
</tr>
<tr>
<td>Resource savings</td>
<td></td>
</tr>
<tr>
<td>Better decision making</td>
<td></td>
</tr>
</tbody>
</table>
Microsoft and Amazon are making big bets on retail IoT.

Microsoft

“Announcing new Microsoft Dynamics 365 AI-driven insights applications and our vision for the future of retail.”

(September 23, 2019)

1. Aims to “help physical retailers understand and improve the in-store experience.
2. Analyze data from video cameras and the internet of things (IoT) sensors, which provide real-time and predictive insights to improve decision making.
3. Improve register experience, monitor refrigerator and freezer temperatures, etc.
How can we navigate this digital transformation in such a way that we realize the positive benefits and avoid the negative outcomes?
Using Data Science to Proactively Manage the Connected C-Store

Difficulties with managing connected and IoT devices.

Using industry analogs to navigate a digital transformation.

Use cases for more effectively managing the connected petroleum retail environment.

Tips for experimenting with predictive maintenance approaches.
Other industries are undergoing similar transitions and can provide helpful analogs for managing complex and connected systems.

- **MANUFACTURING**
  - Predicting device failures (Condition-based maintenance)

- **ENERGY & UTILITIES**
  - Outage management

- **TRANSPORTATION**
  - Advanced Safety Systems
Applying digital transformation lessons from other industries to convenience retail:

- **CONDITIONED-BASED (PREDICTIVE) MAINTENANCE** Identify features that can accurately determine the condition of in-service equipment in order to estimate when maintenance should be performed.

- **OUTAGE MANAGEMENT** Identify features that can help predict causes of outages and poor network performance; effectively model demand in order to help plan ahead.

- **ADVANCED SAFETY SYSTEMS (SELF-HEALING)** Implement backup/secondary processes that will automate responses when threats and issues arise.
The petroleum retail environment is critical to our country’s resilience.

• “When most community residents are evacuating to safety, convenience stores typically remain in operation as long as they can without putting their own employees’ safety and well-being in harm’s way.”

• “Their focus is to see how long can they stay open before the storm, and how they can be first to open when it is safe to do so—all while ensuring that they have power and supply. This means making sure fuel, food and other necessities are available for emergency workers and customers seeking a return to normalcy.”

NACS DECEMBER 2017
There are 16 critical infrastructure sectors whose assets, systems, and networks, whether physical or virtual, are considered so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof.

Convenience stores fall into four categories: Transportation, Emergency Services, Energy, Food and Agriculture, The petroleum retail environment is critical to our country’s resilience.
Using Data Science to Proactively Manage the Connected C-Store

- Difficulties with managing connected and IoT devices.
- Using industry analogs to navigate a digital transformation.
- Use cases for more effectively managing the connected petroleum retail environment.
- Tips for experimenting with predictive maintenance approaches.
CONDITION-BASED MAINTENANCE

Using key signals to prioritize device maintenance and preempt device failures.
How can we forecast device failures and avoid unnecessary and costly device maintenance?
Condition-based maintenance for IoT.

1. When assessing an IoT device, start with its security. Ensure that it is (1) using secure communication protocols, the cloud environment (if it is cloud managed) has been thoroughly assessed, and that the device itself will segmented away from the CDE.

2. Obtain the manufacturer’s device info. Assess the parts to ensure they are up to standard, much as you would for ATGs.

• CMOS battery
• Processor type
• Network Adaptor
• Memory
• Hard drive
Condition-based maintenance - study manufacturer information closely

1. Use data from the manufacturer to potentially set benchmarks and thresholds ahead of time.

- Mean time to Failure (before purchase) and Mean time between failure
- Mean time to repair (total maintenance time/total number of repairs)
1. Based on this information, start putting together an IoT device maintenance plan.
Condition-based maintenance - study manufacturer information closely

1. Measures like MTTF and MTBF are typically ranges (e.g. 2 – 5 years). This begs the question, what factors impact whether a device will fail at the lower end or higher end of the range?

Potential Features

- Device Temperature
- Device CPU/memory utilization
- Ping request responsiveness
- Battery/Power consumption (low battery)
- Other device alarms
- Previous maintenance activity
Condition-based maintenance – develop a plan to collect this data.

1. PCs at the store or another edge computing device, middleware, etc.
2. APIs available?
3. How are these devices networked locally?
4. How frequently should they be polled? 5 minutes, hourly, daily, weekly?
5. Setting thresholds.
6. What tools will analyze the data in the cloud?
Securing Your Data

- Communication Protocols - Use industry best practice
  - Are outdated / insecure protocols in use?
    - SSL
    - Telnet
    - FTP
- Cloud environment assessment
  - What cloud environment are they hosting their servers on?
  - Do they have a vulnerability management program?
  - Do they have compliance attestations?
- Other best practices — segmentation
  - Segment IoT devices from CDE
  - Best practice to have a 3rd Party network segment
Collecting Data

- PC or Edge Computing device
  - Can the Back Office computer be used as a probe?
- Agent vs. Agentless
  - If software can be installed on the endpoint being monitored, typically more information is available
  - If software can’t be installed, network probes can use available protocols to collect information
    - WMI
    - SNMP
    - Syslog
    - APIs (Calls to local devices or to cloud servers)
    - and more!
Use tools to collect data.
Scripts!

```python
from pysnmp.hlapi import *
import sys

def walk(host, oid):
    for (errorIndication, errorStatus, errorIndex, varBinds) in nextCmd(SnmpEngine(),
        CommunityData('public'),
        UdpTransportTarget((host, 161)),
        ContextData(),
        ObjectType(ObjectIdentity(oid))):
        if errorIndication:
            print(errorIndication, file=sys.stderr)
            break
        elif errorStatus:
            print('%s at %s' % (errorStatus.prettyPrint(),
                errorIndex and varBinds[int(errorIndex) - 1][0] or '?'),
                file=sys.stderr)
            break
        else:
            for varBind in varBinds:
                print(varBind)

walk('192.168.1.52', 'iso.3.6.1.2.1.4.7.1.17.1.4.10.9.6.0.24.3.0.0.2.1.4.0.0.0.0')
```
Sending Data to SIEM - Alerts vs Continuous Transfer vs. Reports

- **Alerts**
  - Can be used for real time issues
  - Small amount of network traffic
- **Continuous Transfer**
  - Real time information (not just issues)
  - Consistent network traffic
- **Daily / Weekly results**
  - Benchmarking
  - Performance review
  - Large amount of network traffic, but short amount of time
  - Use collected data to determine schedules for uploading data
What to do with this data?

- Plan Scheduled Maintenance for less congested network times
  - Reduce failure of more “network demanding” tasks:
    - Security Patching
    - Updating in-store / at pump video content
    - Exporting information
    - Vulnerability scans

- Review trends in performance to identify:
  - Failing devices
  - Inefficiencies / areas of improvement
  - Ways to increase uptime
Have process in place to review collected data to better understand network performance

![Network Performance Graph](image-url)
Have process in place to review collected data to better understand network performance
Sending Data to SIEM - Alerts vs Continuous Transfer vs. Reports

- **Alerts**
  - Can be used for real time issues
  - Small amount of traffic on network

- **Continuous Data Transfer**
  - Real time information (not just issues)
  - Consistent traffic on network

- **Daily / Weekly results**
  - Benchmarking
  - Performance review
  - Large amount of traffic on network, but short amount of time
  - Use collected data to determine schedules for uploading data
Condition-based maintenance – it’s a dynamic process.

1. Don’t ignore devices just because they are not failing.
2. If a set of devices fall beyond the median range for memory, CPU, temperature or another feature, ask why?
3. Can anything on that device be adjusted to bring it inline with the others?
4. Don’t have to worry as much about devices that are performing well.
5. APIs available?
6. How are these devices networked locally?
7. How frequently should they be polled? 5 minutes, hourly, daily, weekly?
8. Setting thresholds.
9. What tools will analyze the data in the cloud?
OUTAGE MANAGEMENT

Using a diverse set of features to predict outages and manage overall network risk.
1. Many IT teams make network related decisions using “tribal” knowledge about stores:
   • Stores that are in remote locations
   • Stores that get bad cell signal
   • Stores that are Oil Company A’s network, vs B’s network
   • Stores that have food services
   • Stores that have certain POS systems
Additional information can potentially create a more effective set of groups, especially when it comes to maintaining network infrastructure.

1. Site Information
2. Asset Information
3. Network Performance Data
4. Electric Power Data
Having a single source of truth, featuring a wide set of information, can have a meaningful impact in being able to triage issues quickly between multiple parties.

1. Service Contractors
2. Operations
3. IT
4. POS
5. Network Teams
6. Security Teams
Site Asset Data (according to Conexxus): XML schema provides for reporting of site information (e.g., Site Name, Branding, Location, IDs), as well as devices, including but not limited to:

- POS terminals,
- Payment terminals,
- Scanners,
- Site controllers,
- Fuel controllers,
- Dispensers,
- Carwash controllers
- Tank gauges.

Information about these devices may include the:

- Vendor
- Model
- Versioning (e.g., hardware, software, operating system, firmware and/or kernels)
- Networking configuration
- Device capabilities (e.g., EMV or encryption).
- Device-specific information (e.g., blending capabilities and configuration for a dispenser) is also included in the schema.
Determine CLEAR methods for outage classification following incidents (post mortems)

1. Power Outage?
2. Environmental Related?
3. Human Error?
4. Network Bottlenecks?
5. ISP or Cellular Connectivity?
6. Hardware Failure?
7. Software Bug or Corruption?
Site Information Example

1. Building Age
2. Store Layout
3. Geographic Location
4. Brands, Equipment
5. Square Footage
# Outage Data

<table>
<thead>
<tr>
<th>Ticket #</th>
<th>Start Date</th>
<th>End Date</th>
<th>Start Time</th>
<th>End Time</th>
<th>Store #</th>
<th>Total Duration</th>
<th>Store Traffic</th>
<th>Root Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Power Outage</td>
</tr>
<tr>
<td>109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Human Error</td>
</tr>
<tr>
<td>032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hardware Failure</td>
</tr>
<tr>
<td>056</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Software Bug</td>
</tr>
<tr>
<td>007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Site Asset Data Included

<table>
<thead>
<tr>
<th>Outage Duration</th>
<th>Store</th>
<th>Square Footage</th>
<th>Average Interface Response Time %</th>
<th>Router</th>
<th>Switch</th>
<th>Modem</th>
<th>Cables</th>
<th>ISP</th>
<th>Backup</th>
<th>Geographic al Location</th>
<th>Traffic</th>
<th>Wi-Fi?</th>
<th>Number IP range</th>
<th>Bandwidth</th>
<th>Through put</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 hours</td>
<td>097</td>
<td>2,100</td>
<td>99</td>
<td>SonicWALL</td>
<td>Netgear</td>
<td>Arris BGW120</td>
<td>Cat5</td>
<td>ATT</td>
<td>Verizon MiFi</td>
<td>High</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 hour</td>
<td>109</td>
<td>2,500</td>
<td>99</td>
<td>SonicWALL</td>
<td>Netgear</td>
<td>Arris BGW120</td>
<td>Cat5</td>
<td>ATT</td>
<td>Verizon MiFi</td>
<td>Medium</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 hours</td>
<td>032</td>
<td>2,400</td>
<td>98</td>
<td>SonicWALL</td>
<td>Netgear</td>
<td>Netgear CM500</td>
<td>Cat5</td>
<td>Spectrum</td>
<td>Cradlepoint</td>
<td>Low</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 hour</td>
<td>056</td>
<td>2,400</td>
<td>99</td>
<td>Meraki</td>
<td>Cisco</td>
<td>AC1750</td>
<td>Cat6</td>
<td>Comcast</td>
<td>Verizon MiFi</td>
<td>Medium</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 hours</td>
<td>007</td>
<td>2,300</td>
<td>98</td>
<td>SonicWALL</td>
<td>Netgear</td>
<td>Arris BGW120</td>
<td>Cat5</td>
<td>ATT</td>
<td>Cradlepoint</td>
<td>Low</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Network Data Example

<table>
<thead>
<tr>
<th>Back office Reboots</th>
<th>POS CPU Mean %</th>
<th>BACK OFFICE CPU Mean %</th>
<th>ATG CPU Mean %</th>
<th>POS Memory Mean %</th>
<th>Back Office Memory Mean %</th>
<th>ATG Memory Mean %</th>
<th>ATG Reboots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back office Reboots</td>
<td>POS CPU Mean %</td>
<td>BACK OFFICE CPU Mean %</td>
<td>ATG CPU Mean %</td>
<td>POS Memory Mean %</td>
<td>Back Office Memory Mean %</td>
<td>ATG Memory Mean %</td>
<td>ATG Reboots</td>
</tr>
<tr>
<td>Total POS</td>
<td>Total BO systems</td>
<td>ATG Uptime %</td>
<td>POS Reboots</td>
<td>ATG Reboots</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of Pings Dropped</td>
<td>Packet Loss External</td>
<td>Packet Loss Internal</td>
<td>Average Interface Response Time</td>
<td>POS Uptime %</td>
<td>Back Office Uptime %</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Network Data Example

<table>
<thead>
<tr>
<th>STORE</th>
<th># of IPs</th>
<th># of MAJOR STORE NETWORK OUTAGES S &gt; 20 minutes</th>
<th>PRIMARY ISP MEAN THROUGHPUT (mbps)</th>
<th>CELLULAR MIFI MEAN THROUGHPUT (mbps)</th>
<th>TOTAL WIRELESS FAILOVER EVENTS (ISP drop) &gt; 1 minute</th>
</tr>
</thead>
</table>

**Notes:**
- **STORE** refers to the type of network service provider
- **# of IPs** represents the number of internet protocol addresses
- **# of MAJOR STORE NETWORK OUTAGES S > 20 minutes** indicates the number of network outages lasting more than 20 minutes
- **PRIMARY ISP MEAN THROUGHPUT (mbps)** shows the average throughput of the primary internet service provider in megabits per second (mbps)
- **CELLULAR MIFI MEAN THROUGHPUT (mbps)** displays the average throughput of the cellular MiFi device in mbps
- **TOTAL WIRELESS FAILOVER EVENTS (ISP drop) > 1 minute** refers to the total number of wireless failover events due to ISP drop lasting more than 1 minute
# Power Data

<table>
<thead>
<tr>
<th>Store #</th>
<th>Mean POS Energy Consumption (kWh)</th>
<th>Mean ATG Energy Consumption (kWh)</th>
<th>Square Footage</th>
<th>Surge Protection During Event?</th>
<th>Mean Back Office Energy Consumption (kWh)</th>
<th>Mean Car Wash Energy Consumption (kWh)</th>
<th>Std. POS Energy Consumption (kWh)</th>
<th>Std. Car Wash Energy Consumption (kWh)</th>
<th>Std. ATG Energy Consumption (kWh)</th>
<th>Std. Back Office Energy Consumption (kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>097</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>032</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>056</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
More Effective Methods for Classification.

Using data to forecast store requirements:
1. What stores need multiple WAN connections?
2. What stores need to have a certain wireless backup solution?
   1. ISP drops + MiFi performance
3. What stores need to have which data plan?
   1. Throughput metrics can determine that
4. What stores need to replace power infrastructure?
   1. No surge protection, older stores, UPS’s without battery
Determining Unexpected Relationships

![Box plots and scatter plots showing relationships between sepal and petal lengths for three species: setosa, versicolor, and virginica.](image)
Finding Inspiration

• A prediction is a definitive and specific statement about when and where an earthquake will strike: a major earthquake will hit Kyoto, Japan, on June 28.
• Whereas a forecast is a probabilistic statement, usually over a longer time scale: there is a 60 percent chance of an earthquake in Southern California over the next thirty years.

Nate Silver, The Signal and the Noise
“Small Data” can still help to improve probability.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Prediction</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Algorithm A gives a 2% increase in true positive detection of threat X over Algorithm B”</td>
<td>“We are targeting 80% of all servers of class Y to be have a security grade of at least A in the next 3 months”</td>
</tr>
<tr>
<td>Methodology</td>
<td>Use knowledge of potential threats to <strong>hypothesize</strong> security improvements</td>
<td>Combine business goals, gut feeling + data from multiple sources to set a <strong>goal</strong> for security performance</td>
</tr>
<tr>
<td>Guiding Question</td>
<td>What value should I assign this algorithm?</td>
<td>Given what we know about our threat model(s), how do we plan the path forward?</td>
</tr>
</tbody>
</table>
Determining Unexpected Relationships

Site and Asset Relationship Examples
Connectivity Failures vs Backroom Temperature?
Connectivity Outages vs Building Age?
Connectivity Outages vs Food Vendor/QSR (yes or no)?
Connectivity Failures vs Square Footage?
Connectivity Outages vs Surge Protection State?

Network Relationship Example
Connectivity Failures vs Dropped Packets?
Connectivity Outages vs POS Reboots in the last year?
Connectivity Outages vs # VPN tunnels?
Connectivity Outages vs # IPs?
Connectivity Outages vs Throughput?
SELF-HEALING

Using simple, rule-based automation to implement self-correcting mechanisms on systems.
Self-healing POS application failures – a true story.

1. POS services not starting after reboots following application of weekly patches.
Set a baseline – measure the impact of the failure.

1. Number of stores impacted?
   • 40/100 stores

2. How long until the failure was noticed? Why not sooner?
   • 1 hour.
   • Absolutely critical or not-critical?

3. Duration of impact?
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3202server, Windows Service - POS Integration Manager 3 Service, Failed, Restart Service-POS Integration Manager 3 Service, Failed, Trigger State = Failed</td>
</tr>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3202server, Resolution: Success, Service: Windows Service - POS Site Controller SSDPMonitor, Restart Service-POS Site Controller SSDPMonitor, Trigger State = Failed</td>
</tr>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3202server, Resolution: Success, Service: Windows Service - POS CMC Bridge, Restart Service-POS CMC Bridge, Trigger State = Failed</td>
</tr>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3202server, Resolution: Success, Service: Windows Service - POS API Router, Restart Service-POS API Router, Trigger State = Failed</td>
</tr>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3147server, Windows Service - POS Services, Restart Service-POS Services, Trigger State = Failed</td>
</tr>
<tr>
<td>9/22/19 5:21:00 AM</td>
<td>Device: 3147server, Resolution: Failed, Service: Windows Service - POS Integration Manager 3 Service, Restart Service-POS Integration Manager 3 Service, Trigger State = Failed</td>
</tr>
</tbody>
</table>
Measure What Matters
Measure What Matters

~ 25% - 30% of total endpoints are having POS services stop on Wednesdays
Develop hypotheses

1. Find a machine that had an issue and did not have an issue.
2. Look through events of both machines to identify which event happened on the machine that had the issue vs which did not have the issue.
3. Event was identified alongside additional dependent events.
4. Verified that the other failed machines also had the event.
5. To this point, still did not WHY the service was not starting. To this point, only knew that the machines were not starting post patches. Root cause analysis not yet done.
Develop hypotheses

HYPOTHESIS A – Resource consumption impacting ability for SQL service to start.
• Sample set of reboots...trying to get somewhere close to 40%, which then requires a sufficient set of machines. Went ahead – with customer’s permission – and scheduled a set of reboots even though patches were not applied.

HYPOTHESIS B – Network-related
• Misconfiguration in SQL/POS software
• Issues related to the recently applied Windows patch.
Root cause analysis – try to determine the critical relationship, either via correlation or domain-knowledge about dependencies.

1. Identified that reboots were the problem triggering the application failure.
2. Unfortunately, root cause is inherent in the system – Windows 7 does not provide options for priority on boot.
   • Automatic discovery and correction of faults.
Implementing Self-Healing Actions

1. A/B tested different self-healing methods on a subset of machines following schedule reboots after patches to compare effectiveness of script.
2. Important to be granted to freedom of carrying out an experimental approach. Requires management support.
3. Continue ongoing performance monitoring.
Self Healing Actions

Flowchart:
- Start
- Is POS service running?
  - Yes: Start POS Service
  - No: Start POS Service
- End
Self Healing Actions

Start

- Is POS service running?
  - Yes: End
  - No: Start POS Service

- Is SQL Server service running?
  - Yes: Wait 5 minutes
  - No: Start SQL Server service

Start POS Service

Start SQL Server service
Lessons Learned – same approach can be followed for a variety of anomalous system failures that do not have a direct fix.

1. Rule-based power cycling using power distribution units (PDU), in the event of failed ping tests.
3. So long as you are monitoring the right things and understand the dependencies, and you have the ability to remotely manage/remotely impact the device, then you can implement these rules in your enterprise.
Using Data Science to Proactively Manage the Connected C-Store

- Difficulties with managing connected and IoT devices.
- Using industry analogs to navigate a digital transformation.
- Use cases for more effectively managing the connected petroleum retail environment.
- Tips for experimenting with predictive maintenance approaches.
Ask questions. Then ask more questions.
Ask questions. Then ask more questions.
Develop minimal viable products to determine the potential impact of data products/strategies.
Monitoring and analysis tools used to be defined in separate product categories.

<table>
<thead>
<tr>
<th></th>
<th>MONITORING</th>
<th>ANALYSIS</th>
<th>COMMUNICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERNAL</td>
<td></td>
<td><strong>splunk</strong></td>
<td><strong>pagerduty</strong></td>
</tr>
<tr>
<td>CUSTOMER</td>
<td></td>
<td><strong>zendesk</strong></td>
<td></td>
</tr>
</tbody>
</table>
Not the case today – application categories are merging.
Explore new approaches to building your organization’s technology/application stack.

<table>
<thead>
<tr>
<th>INTERNAL</th>
<th>CUSTOMER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MONITORING</strong></td>
<td><strong>ANALYSIS</strong></td>
</tr>
</tbody>
</table>

"Dev ops," traditional monitoring, IT Service Management (ITSM), IT Operations Management (ITOM), Asset management, Alerting, Security Information and Event Management, Data Analysis/Data Science, and CRM tools have all crossed boundaries.
Leverage APIs to ensure that applications play well together.

<table>
<thead>
<tr>
<th></th>
<th>MONITORING</th>
<th>ANAYLSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERNAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUSTOMER</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Dev ops,” traditional monitoring, IT Service Management (ITSM), IT Operations Management (ITOM), Asset management, Alerting, Security Information and Event Management, Data Analysis/Data Science, and CRM tools have all crossed boundaries.

CONEXXUS  solve forward
Communities and internal forums centered around certain tools (e.g. SIEM, data analytics, network monitoring) are great sources of knowledge, especially for answering questions like “what can I measure,” ”how can I measure it,” and “what insights can be made from the data?”
Leverage communities and events to find analogs that apply to your organization.

1. Data science meetups and Kaggle – great for seeing examples of data science being used in a wide variety of industry applications. Especially good for getting ideas on how to select features, ask questions, and structure data.

2. Pay attention to site reliability engineering (SRE) and DevOps communities – forums, tools, code, etc.

3. Talk to engineers from other analogous industries (energy & utilities, manufacturing, transportation, healthcare) to see how they think about ensuring availability and high performance of critical systems and devices.

4. Keep talking to vendors/manufacturers – know experts who can provide proper consultation about devices – their level of security, reliability, ease of management, etc.
Using Data Science to Proactively Manage the Connected C-Store

- Difficulties with managing connected and IoT devices.

- Using industry analogs to navigate a digital transformation.

- Use cases for more effectively managing the connected petroleum retail environment.

- Tips for experimenting with predictive maintenance approaches.
Key Takeaways

1. It is worth implementing new connected/IoT devices if the business case makes sense, the devices abide by the proper security standards, and the overall enterprise and operations is prepared to effectively manage the device following implementation.
2. Preparation includes ensuring that a wide array of data points are being collected, tracked, and made visible to the appropriate parties.
3. Preparation also includes having the right operations and monitoring tools in place to ensure that the right data can be collected.
4. Always have regular retrospective/post mortem meetings to ensure that any incidents/outage events are properly classified. You cannot collect data only when it’s necessary; must be done long before.
5. Regularly perform exploratory data analysis using different variables to see whether any interesting relationships may exist that can help prevent failure or outages downstream.
Speakers

Ashwin Swamy
Director of Cyber Resilience, Omega ATC
ashwin.swamy@omegaatc.com

Thomas Duncan
Senior Security Strategist, Omega ATC
thomas.duncan@omegaatc.com
Q & A
• Website: www.conexxus.org
• Email: info@conexxus.org
• LinkedIn Profile: Conexxus.org
• Follow us on Twitter: @Conexxxusonline
DISCLAIMER: Conexxus does not endorse any products or services that may be described or mentioned in this presentation. The views and opinions expressed in this presentation are solely those of the speakers and not of Conexxus. By hosting this webinar, Conexxus is not providing any legal advice; if you have any questions about legal issues raised or discussed, you should seek the assistance of attorneys who are competent in that area.