Copyright © 2015 Splunk Inc.

.conf2015

Deeper Insights into Water Treatment (and everything else) through Splunk

Andy Kuhn

Application Developer, DevOps

(Certified Splunk Architect)

Denver Water

splunk>

Agenda

- Introduction
- What is Denver Water?
- Adoption and Growth of Splunk
- Customer Self-Service
- Operational Intelligence
- Water Treatment and Deep Data Analysis – where we are now and where we are going
- The Data Rub
- Future uses

Introduction

- Andrew Karl Kuhn, call me 'Andy'. Spent most of my life in Colorado and New Mexico.
- Degree in Chemistry, worked professionally as a biochemist for 6 years. Been working in IT since 2000 primarily in private industry.
- Application development, architecture, middleware, devops
- Music performance and recording, ancient history, cars, backpacking, anything educational, home maintenance

What is Denver Water?

- **Public agency established in 1918**, whose revenues are derived from water rates and tap fees (not taxes), serving about 1.3 million in Denver and the surrounding area
- **~1000 employees**. Builds and maintains all collection, transmission, treatment and distribution serving the Denver Metropolitan Area
- **3rd largest land owner** in the state
- **234,000 acre feet (289 million cubic meters)** of water harvested each year
- **2% of all water in Colorado is processed** by Denver Water

Does a Water Utility collect data?

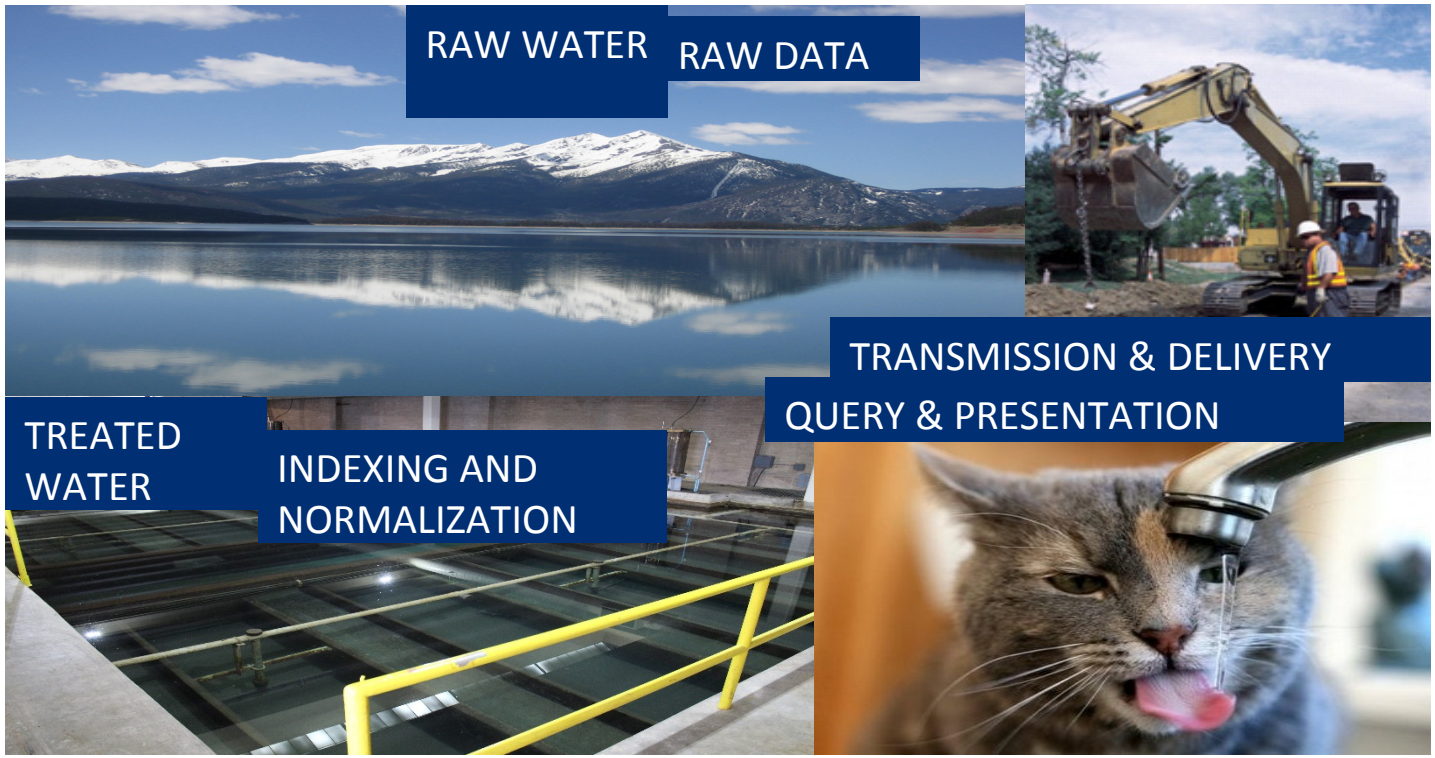
- **Reservoirs, Streams, Weather** (aka Raw Water, non-potable)
- **Treatment Plants** (aka Treated Water, potable)
- **Distribution pipes, compression stations, household taps**
- **IT Enterprise events** (monitoring applications and systems)
- **Desktop management** (monitoring Windows logs for user support)
- **IT Integration activities** (monitoring service execution and performance)
- **IT Security activities** (DMZ access, Active Directory issues)

splunk> @ DENVER WATER

- **From 50 to 1400 Splunk Universal Forwarders** deployed from January to December 2014 (28x increase)
- **Cumulative daily data indexing reduction of 30%** due to careful analysis of indexed information and determination of value, even with the addition of a great deal of additional data.
- **Increase of splunk enterprise utilization** - indexes, sourcetypes, data models and search head usage by 500%
- **Using ~50% of our 50GB license** each day (tiny, but important)
- Beware of unneeded Windows Security Events from Domain Controllers

DENVER WATER

How water is like data



splunk> @ DENVER WATER

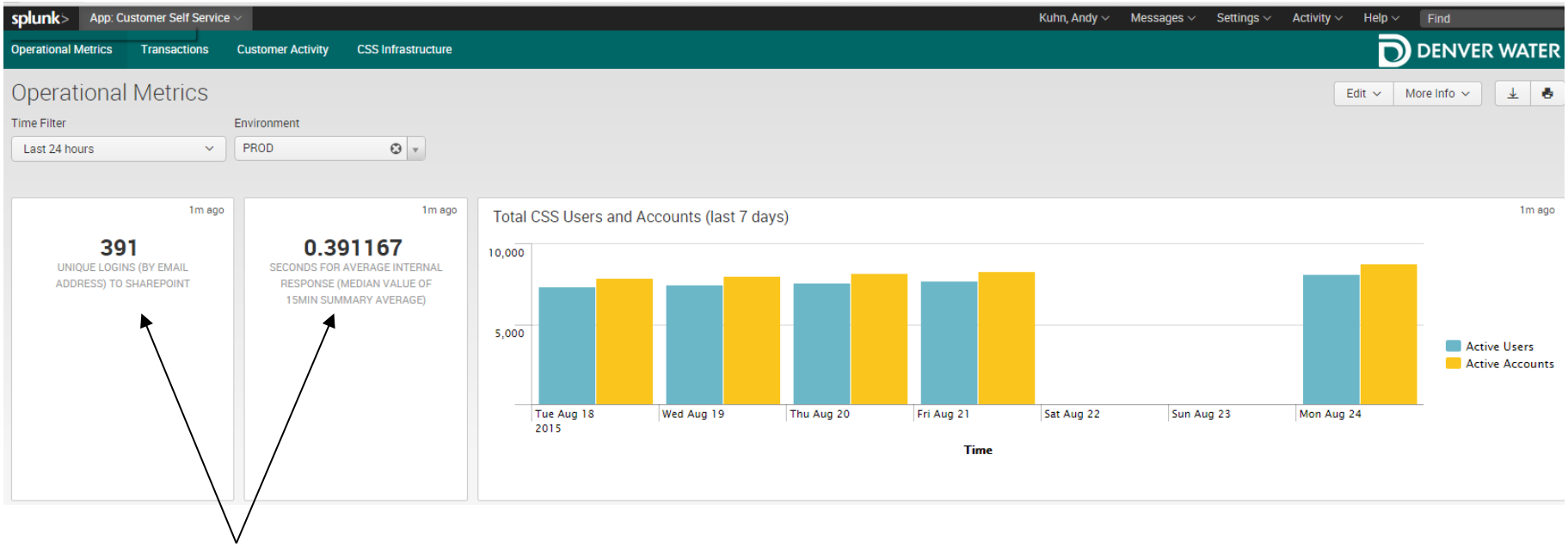
Customer Self-Service Operational Intelligence Water Treatment

DENVER WATER

Customer Self-Service Website Metrics

- **Stakeholders are wide-ranging.** IT developers, Project Managers, Customer Support, Marketing, Public Affairs and Executives
- **Intended to provide comprehensive information** pertaining to site usage as well as individual transaction information
- **Large challenge to gather requirements from and educate a diverse group of stakeholders** on the use of the CSS Splunk dashboards.
- **Summary indexing necessary** for some data to accommodate performance requirements (from last hour to a year ago)

Customer Self-Service Website Metrics



Summarized high-level metrics for managers and executives uses ***tstats*** command.
Precise, clear labelling of data elements to reduce confusion and incorrect assumption

Customer Self-Service Website Metrics

Most Recent CCB Transactions by Person (*)

2m ago

Time	Message	Action	changed values	PersonId	AccountId
08/25/2015 13:25:40.252		Email Address Changes	[redacted]	[redacted]	
08/25/2015 13:22:47.52		Ebill Subscription Changes	flag=Subscribed email=[redacted]	[redacted]	[redacted]
08/25/2015 13:22:35.765		Ebill Subscription Changes	flag=Subscribed email=[redacted]	[redacted]	[redacted]
08/25/2015 13:21:38.524		Phone Number Changes	type=CELL number=[redacted]	[redacted]	
08/25/2015 13:21:02.643		Notification Changes	NOTIFY OUTAGE	[redacted]	
08/25/2015 13:20:54.114		Email Address Changes	[redacted]	[redacted]	
08/25/2015 13:20:44.116		Email Address Changes	[redacted]	[redacted]	
08/25/2015 13:20:40.384		Phone Number Changes	type=CELL number=[redacted]	[redacted]	
08/25/2015 13:20:19.862		Mailing Address Changes	address=[redacted]	[redacted]	[redacted]
08/25/2015 13:19:46.644		Customer Name Changes	[redacted]	[redacted]	

Row_highlighting.js turns all rows (with errors) red with white lettering so easy to spot

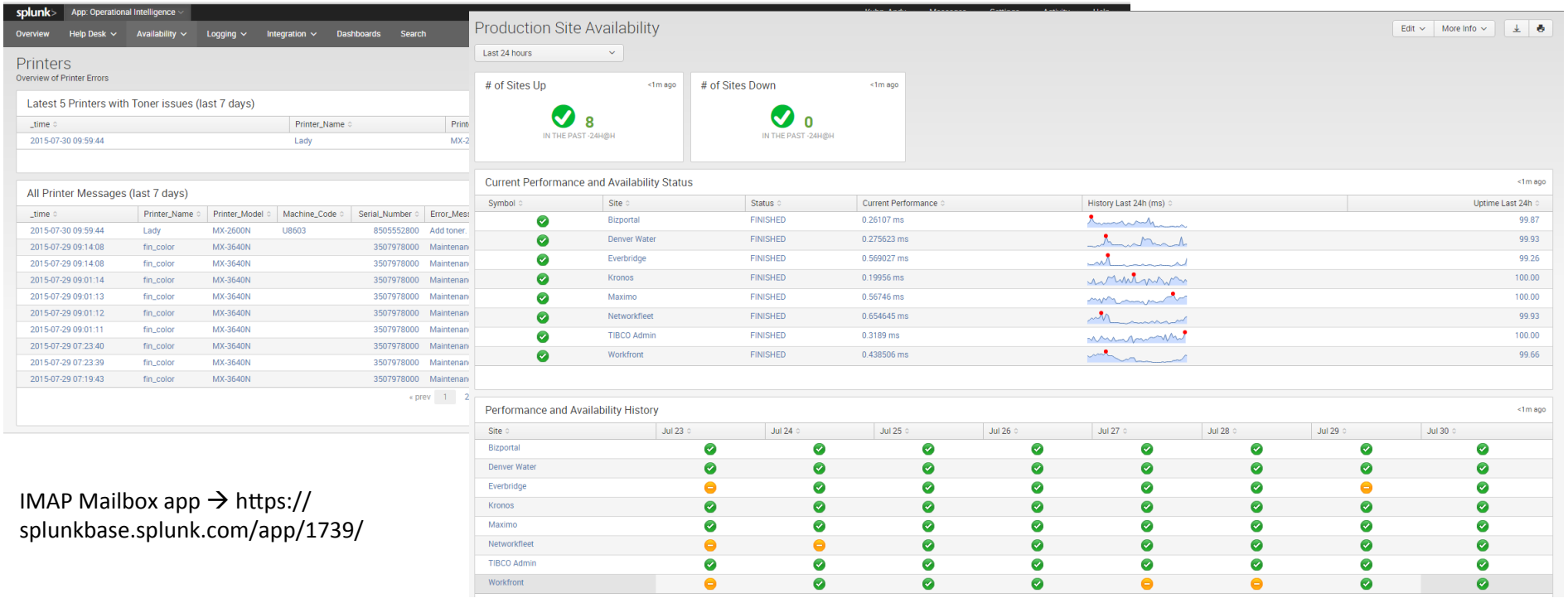
Actual values changed by the customer from the log file. Can search all customers(*) or a specific customer

DENVER WATER

Operational Intelligence

- **Provides critical information** to Help Desk and Desktop Support personnel related to desktop health and printer errors
- **Provides application service and availability monitoring** through scheduled service calls, for internal and externally-hosted applications
- **Provides a means to troubleshoot transactional/integration activities** in any environment through a log viewer that displays all logs based on a variety of selectors and provides the ability to decode and view the payload for any message. Heavily used by development staff

Operational Intelligence



IMAP Mailbox app → https://splunkbase.splunk.com/app/1739/

Operational Intelligence

splunk> App: Operational Intelligence | Kuhn, Andy | Messages | Settings | Activity | Hel

Overview | Help Desk | Availability | Logging | Integration | Dashboards | Search | DENVER WATE

Logviewer

Log viewer for application and integration logs

Edit | More Info

Select Time Range:	Keyword Search	Environment	Component	Severity	Event Type	
Last 15 minutes	*	× PROD	× WAM	× All	× All	Submit

Log Results

1m ag

i	Event_ID	Event Time	Severity	Type	Component	Process Name	Sourcetype	Business Key	Reference ID	Message
	ID:EMS- [redacted]	08/05/2015	Info	Stop	WAM	Services/DWWOPublisher/Processes/DWWOPublisher.process [redacted]	tibco_cbe	15- [redacted]	OutboundTdUpdateWorkItem- [redacted]	Completed proces WorkOrder create/ request and sent r

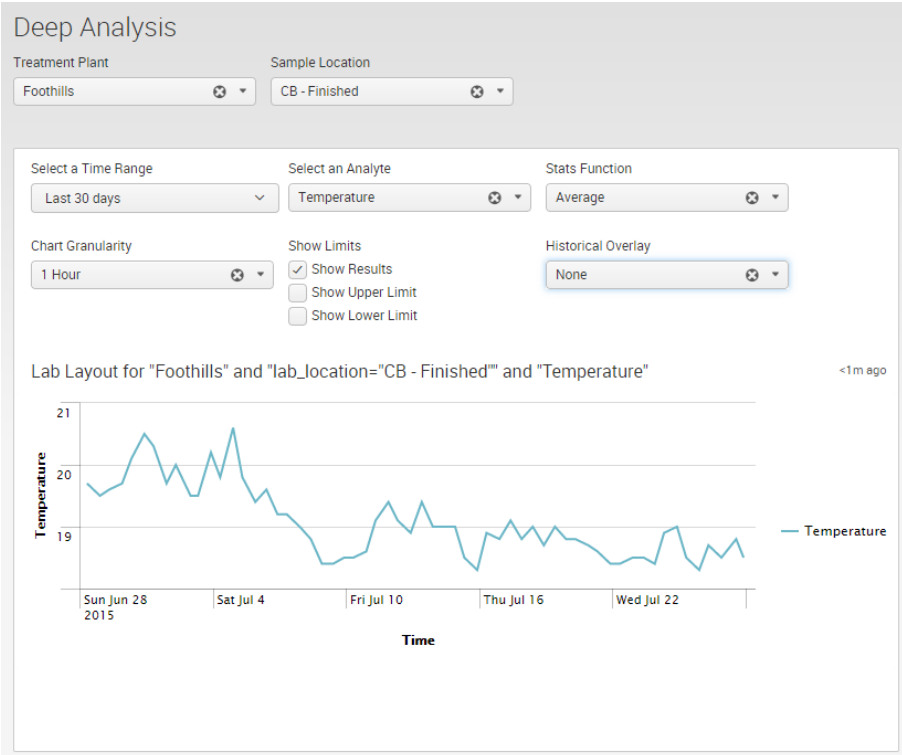
Type	Information
Stack Trace	empty stack_trace
Process Stack	
Payload	<?xml version="1.0" encoding="UTF-8"?> <ns0:ProcessDWEnterpriseWorkOrder xmlns:ns0= [redacted] ader><ns0:Sender><ns0:CorrelationID>OutboundTdUpdateWorkItem-4efd3af4</ns0:CorrelationID> <ns0:SystemID>TIBCO</ns0:SystemID><ns0:ComponentID>DWWOPublisher</ns0:ComponentID><ns0:LogicalID/><ns0:UserID>GAT</ns0:UserID></ns0:Sender><ns0:CreationDateTime>2015-08-05T15:17:33.2135119-06:00</ns0:CreationDateTime> <ns0:MessageID>OutboundTdUpdateWorkItem-4efd3af4</ns0:MessageID><ns0:TaskID/><ns0:MessagingPattern/><ns0:Status>Success</ns0:Status><ns0:ProcessingMessage message="The WorkOrder has been created successfully" type="Success"/> <ns0:Operation>REPLACE</ns0:Operation></ns0:DWSOAHeader><ns0:DWEnterpriseWorkOrder><[redacted]TUS/><ns0:STATUSDATE>2015-08-05T13:30:56-06:00</ns0:STATUSDATE>

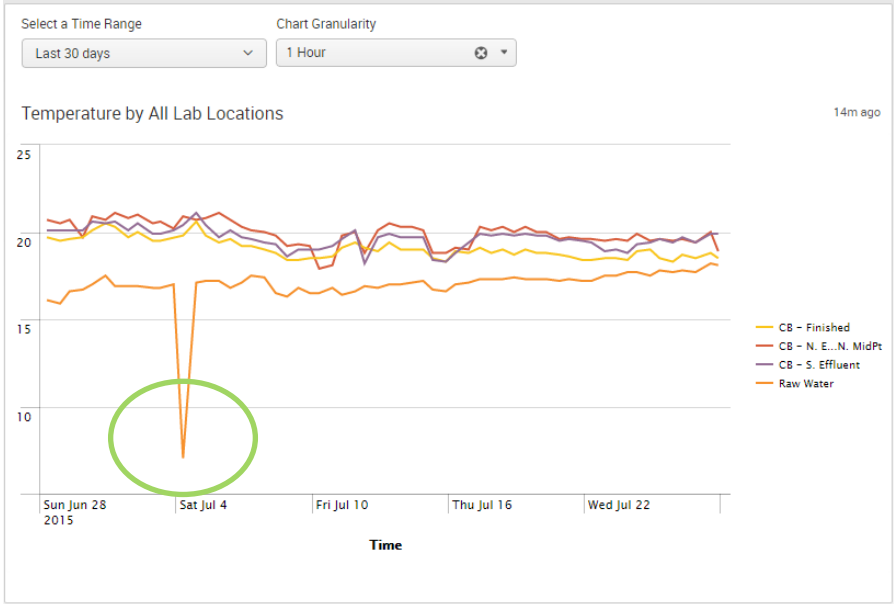
DENVER WATER

Water Quality

- **At least 5 different types of measurement sources** (Raw, Treated, Weather, Labs)
- **Using REST TA and DBX** from our Splunk heavy forwarder as well standard file tail monitoring
- **~20 different Splunk sourcetypes**
- **3 Splunk data models**
- **4 Splunk indexes** with frozen archives, 5 summary indexes
- **Engagement with stakeholders** to understand needs, hopes and fears.

Water Quality – Deep Analysis

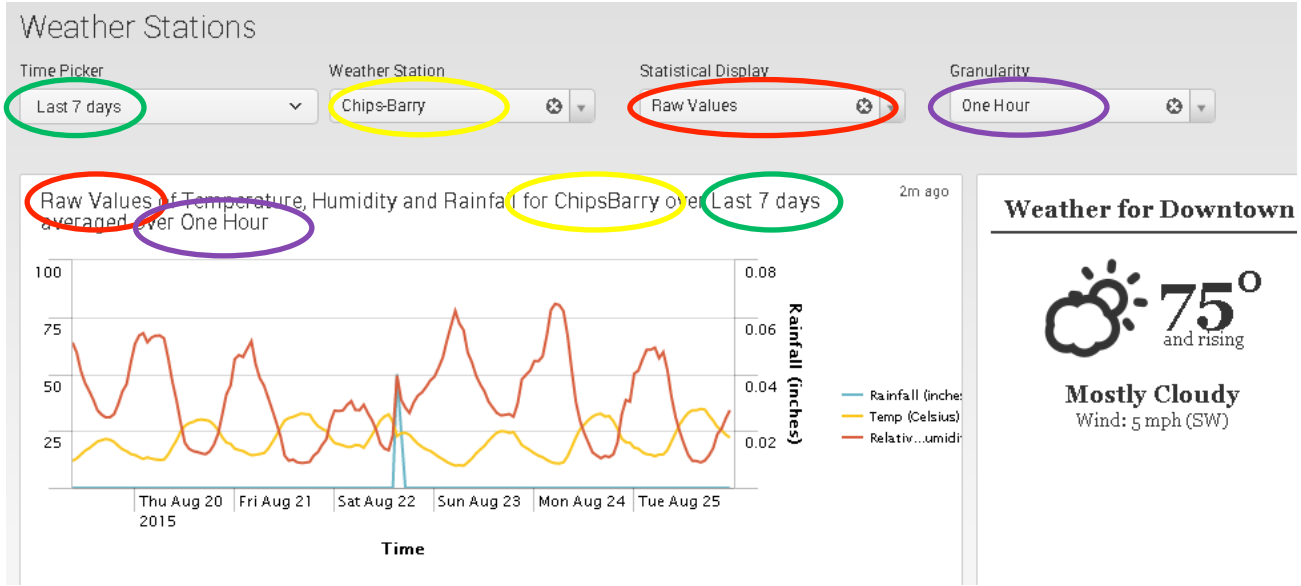




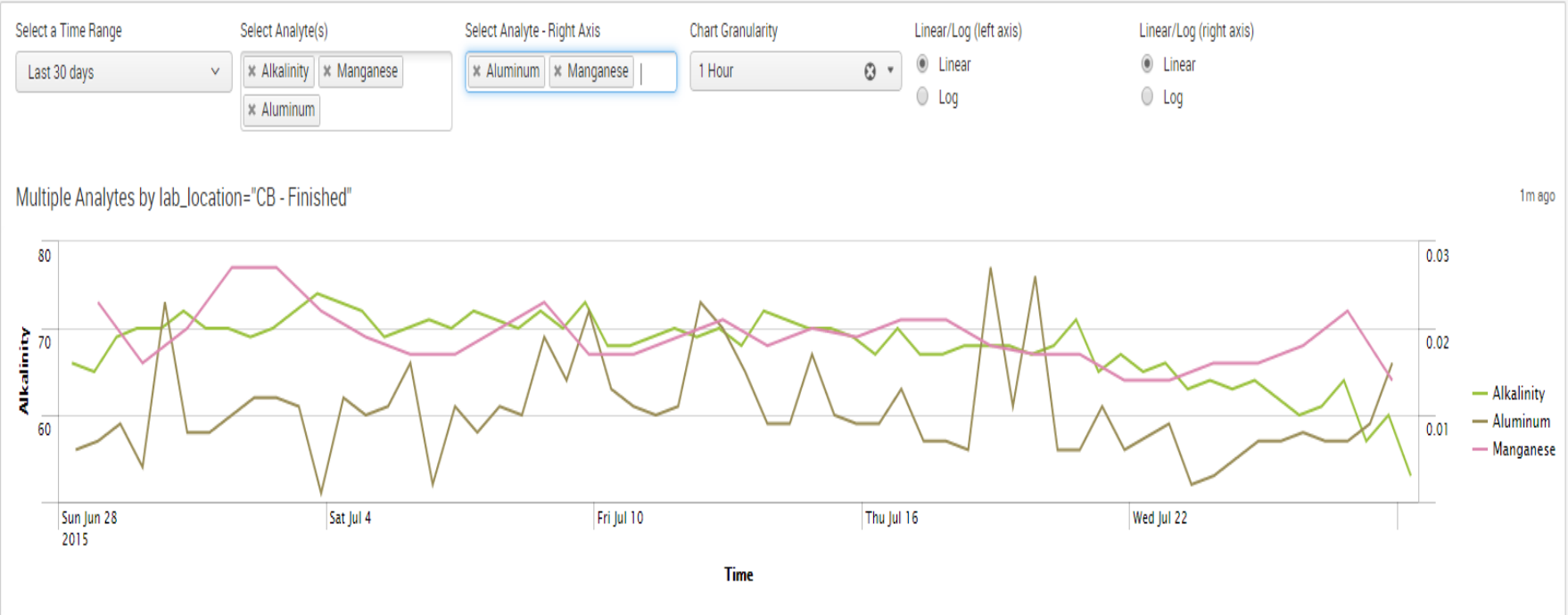
DENVER WATER

Water Quality – 6.2.x dynamic labels

```
<input type="dropdown"
token="weather_station_token"
searchWhenChanged="true">
    <label>Weather Station</label>
    <choice
value="ChipsBarry">Chips-Barry</
choice>
    <choice
value="Kendrick">Kendrick</choice>
    <default>Kendrick</default>
    <change>
     <set
token="station_label_token">$label
$</set>
    </change>
  </input>
```



Water Quality – Qualitative Correlation



DENVER WATER

Water Quality

Lab Sheets

Edit | More Info

Treatment Plant: Foothills

Sample Location: × Raw Water

Analyte: × All

Sample Date: Yesterday

Limit Definition: Original / Stddev2 / Stddev3

Lab Sheet for "Foothills"

1m ago

Location	Analyte	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Lower_Limit	Upper_Limit	Late_Entry	Late_Update
Raw Water	Alkalinity								63												64					20	86	false	false
Raw Water	Hardness																				80					35	150	false	false
Raw Water	Manganese Dissolved																						0.012			0	0.5	false	false
Raw Water	Temperature														18.2								18.1			0	20	false	false
Raw Water	Turbidity				3.84				3.86				3.49				3.2				3.08				3.27	0.4	200	false	false
Raw Water	pH														8.12								7.93			6.7	8.9	false	false

DENVER WATER

Water Quality – The Future

- **Onboard remaining relevant data** associated with the collection and treatment processes
- **Determine additional groups requiring access** to data (conservation, planning, others?)
- **Determine Role Authorizations** for different data sets
- **Specific conversations to further tune dashboards** with managers, staff water chemists, treatment specialists, et al.
- **Setup alerting** for particular conditions as identified in the specific conversations with stakeholders

DENVER WATER

The Data Rub (when change hurts)

- **Conflicts over data ownership** and who should be provided access
- **Confusion in technology ownership and process** has hindered our ability to onboard data (4 completely different ways to collect and store water measurements)
- **Determination of the Source of Record** has not been straightforward. Many people have many different opinions about where data should be written from the generating system and how such data should be handled.
- **Technological differences** between products often dictate how we get the data into Splunk
- **Competing or overlapping technologies at Denver Water**

Question and Answer

Andy Kuhn
Denver Water
Andy.Kuhn@denverwater.org
Andy_Kuhn@yahoo.com
Cell: 303-641-1574


.conf2015
THANK YOU
splunk>