Using Open-Source Tools to Solve Customer Problems Before They Notice

A Technical Paper prepared for SCTE/ISBE by

Andrew Gray
IP Engineer IV
Cox Communications
6305 Peachtree-Dunwoody Rd, B06-151X
+1 (404) 269-9298
Andrew.Gray@cox.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

Introduction

Many RF impairments in the plant are customer impacting, can be difficult to isolate, and in some cases are problems that have slowly built up over time. While a variety of tools exist to troubleshoot and diagnose the problems once they are known, the DOCSIS platform provides monitoring capabilities to allow a constant diagnosis of all nodes at the same time. This paper and its accompanying Cable-Tec Expo technical workshop presentation discuss a system deployed as an experiment in part of the Cox Communications network that uses a variety of open-source software (including Cacti, PHP, and other scripting tools) to automatically monitor, graph, and most importantly, email a daily and weekly summary of the nodes with the most issues. These tools allow a presentation of the parts of the plant down to the individual node level that need the most preventive (and sometimes corrective) maintenance, and provide time-of-day based information that is useful for tracking down issues that repeatedly occur at the same time every day. This system has been recognized as measurably helping to reduce truck rolls and customer call-ins in the region and helping to raise customer satisfaction.

Problem Description

Several years ago, a chance conversation occurred between the author of this paper and a manager of Cox's field services group in Las Vegas. The topic of discussion focused on long-standing RF impairments that continued to exist in the plant, and why it was taking so long to resolve the issues. Cox possessed, even back then, a multitude of tools for troubleshooting issues, for looking at the plant as a whole, and numerous procedures for tracing down problems.

The feedback from the manager was interesting. While all the tools existed in some form, they all suffered from one or more faults that made them tricky to use in the field:

1. They required a full web browser to work, which meant personnel out in the field could not use the variety of smart phones they had with them, instead needing to boot up corporate laptops, which was slow. Applications requiring a full implementation of Java for their user interfaces were usually the culprit, with applications requiring older versions of Internet Explorer making up the rest of the examples.
2. They required multiple levels of authentication to get to, and permissions were routinely lost or not always granted to new personnel coming onboard.
3. They were slow to refresh – some tools were as slow as daily, and none was faster than hourly.
4. They were difficult to use – requiring multiple operations to get to the needed data.
5. They weren't automated enough – while the data was there, no analysis was done to try and automatically filter out the good data from the bad, or highlight problems.

The conversation turned to what elements would be the most useful, which led to the creation of the system documented in this paper.

System Overview

The system created leveraged several existing open source tools to avoid recreating the wheel:

1. Cacti (http://www.cacti.net/) – A system for polling SNMP (simple network management protocol) -based data and storing it in relatively simple-to-use data files.

2. RRDTool (http://www.rrdtool.org/) – Used by Cacti as its on-disk database format
3. PHP (http://www.php.net/) – Used by both Cacti and additional scripts that were already in use, and was leveraged again for this project.
4. PostgreSQL (https://www.postgresql.org/) – A database system, used to store information such as node to service group mappings.
5. Debian Linux (https://www.debian.org/) – The operating system all the tools ran on, this distribution had been used extensively throughout the market for various tasks, a local package repository was available, and a number of engineers could support it.

The system has two main modules – the glue between the real world and Cacti, and a script to handle the reporting aspect.

1. Cacti Interface

Cacti has a built-in integrated high-speed SNMP polling engine referred to as Spine. Cacti also has a simple, but very functional, user interface for viewing the data:

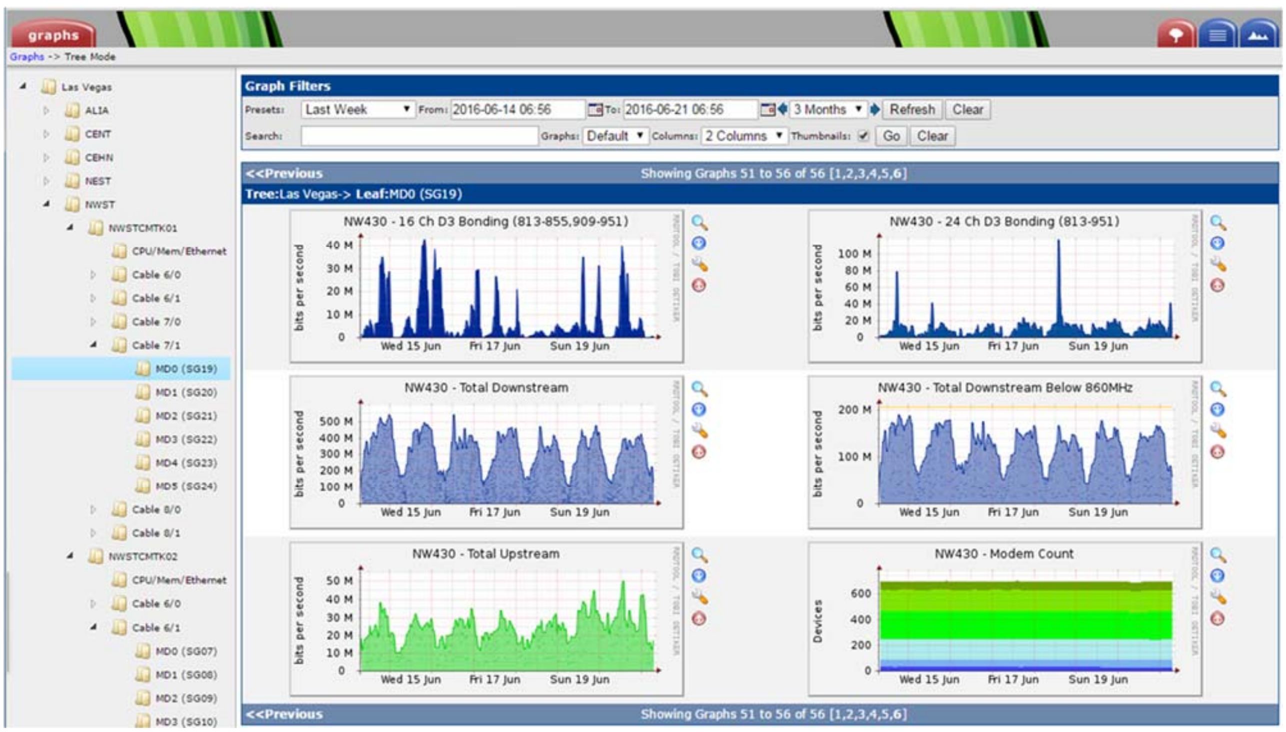

Figure 1 – Main Cacti Layout

A script was created to take the existing node database, stored in PostgreSQL, and create all of the entries inside of Cacti, sorted by market, then hub site, then CMTS, then card, then finally service group, as can be seen in Figure 1. The script uses calls to Cacti's interfaces to register new devices as needed, create the entire hierarchy, and can handle any number of moves, adds, or changes, including doing a full rebuild of the topology if need be.

A number of data sources are created and polled for every service group, including:

1. Traffic, MER (modulation error ratio), FEC (forward error correction) percentage, and FEC counts for every upstream
2. Traffic for every downstream
3. Total traffic upstream, total traffic downstream, and total traffic downstream for QAM carriers below 860 MHz.
4. Modem counts registered on the node

Figures 2-11 give examples of the visualization of the data obtained for each service group. These charts were designed to be very clear and explicit as to what they were showing. Each graph is labeled fully with the node and what information is presented, which permits the graphics to be inserted inline into emails between personnel working on the issue and not need additional explanation.

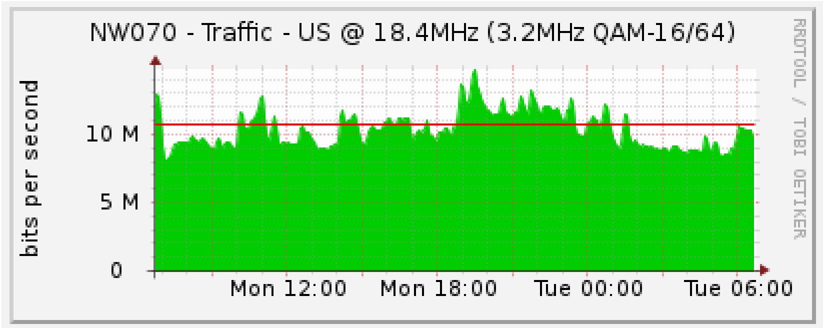

Figure 2 – Upstream traffic chart for a channel

Note the upstream traffic chart includes a red line denoting 70% of channel capacity.

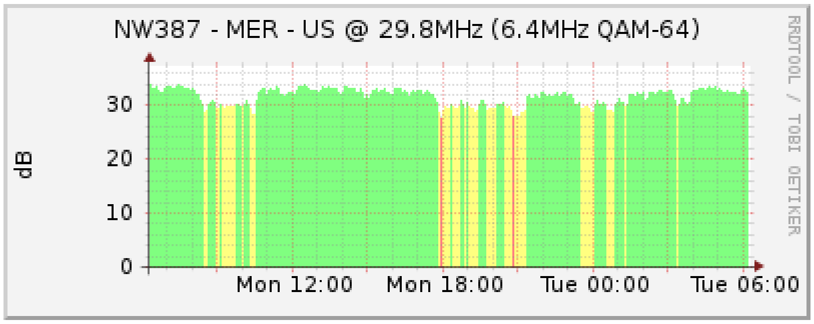

Figure 3 – Upstream MER chart for a channel

The upstream MER chart is arranged so that values over 31 dB are green, 28 dB to 31 dB are yellow, and lower than 28 dB show red.

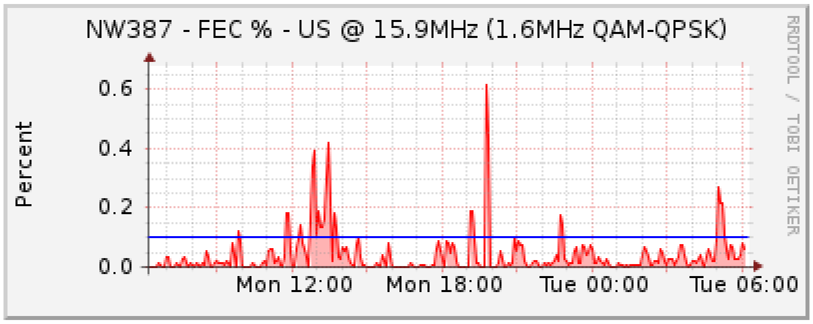

Figure 4 – Upstream FEC percentages for a channel

The FEC percentages chart includes a blue line at 0.1%, denoting the limit where significant performance problems start to occur.

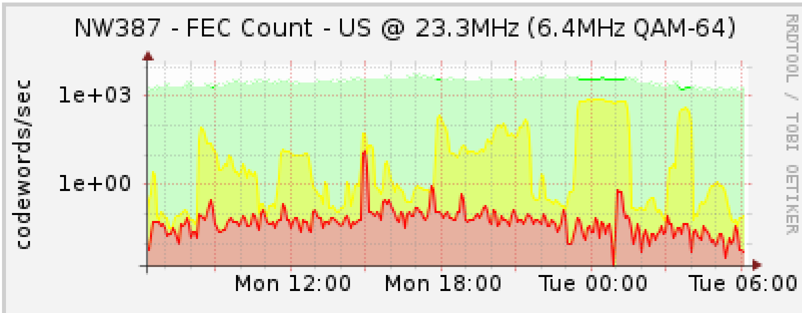

Figure 5 – Upstream FEC counts for a channel

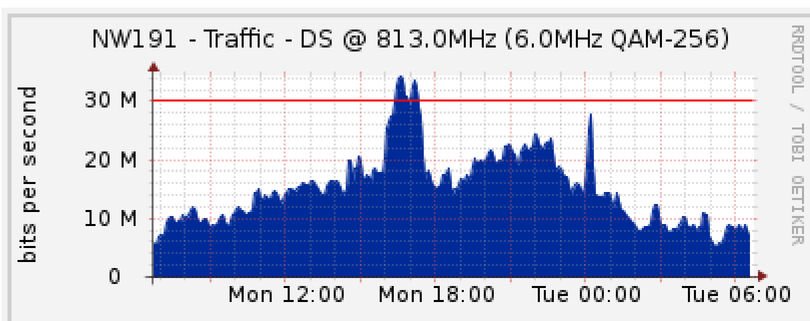

Figure 6 – Downstream traffic for a channel

Note the downstream traffic chart also includes a red line at the 80% saturation point.

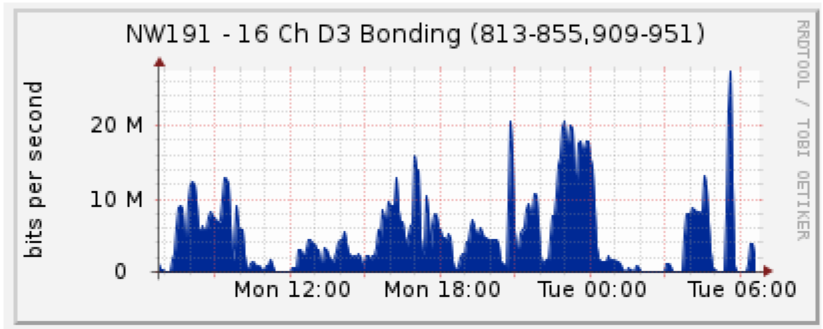

Figure 7 – Downstream traffic for a bonding group

A similar chart is produced for each bonding group active for the given node.

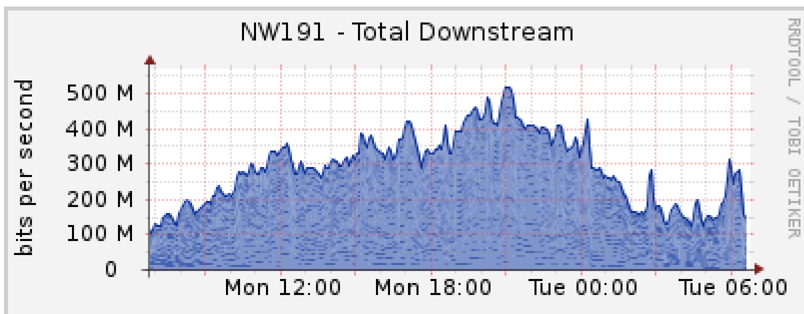

Figure 8 – Downstream traffic for a node

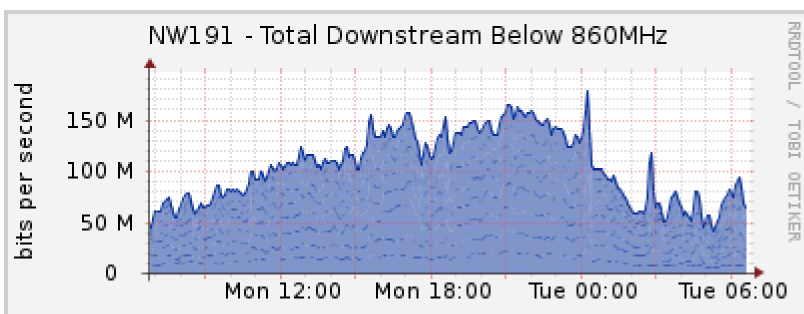

Figure 9 – Downstream traffic for a node, only carriers below 860 MHz

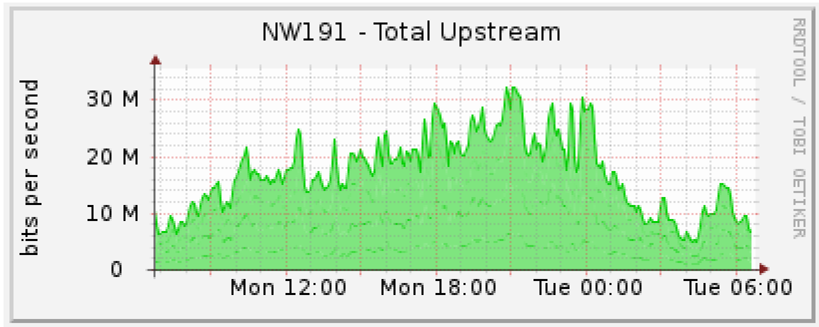

Figure 10 – Upstream traffic for a node

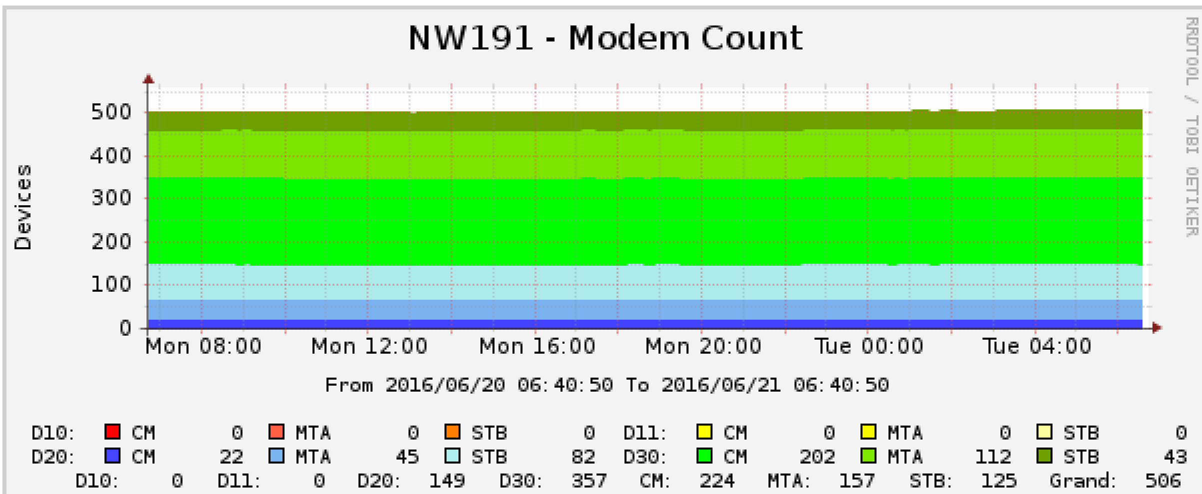

Figure 11 – Modem Counts for the node

The modem count chart was initially added just for curiosity; however it later turned out to be a rather useful chart, especially for its variants that report at the CMTS, hub site, and market levels that total up all the counts beneath them. This data was used in various other reports to judge D3.0 penetration, verify and complete D1.0 and D1.1 device removals, and track the deployment of D3.0 set top boxes (STBs) and multimedia terminal adapters (MTAs)

All of the charts, while shown here zoomed in on a 24-hour period, can be further zoomed in on demand, or zoomed out to show up to a year's worth of statistics. This data is updated every five minutes by Cacti (resulting in almost 800,000 data points being collected by this system covering two markets in every interval). The hardware needs to support this rate of polling are quite minimal – one server is running the polling engine as a virtual machine (VM), along with a number of other functions. These 800,000 data points take on average 70 seconds to collect during each polling cycle.

2. The Reporting

As mentioned previously, other systems were already collecting data in other ways and means, however the Cacti system is a far simpler system for the end user to utilize. Leveraging the data collected by Cacti allowed a number of different automated reports to be written.

The first report is a daily report, emailed to the field technicians, and their managers, containing an automated analysis of all the data collected above on a node by node basis.

Node NW852	**NWSTCMTK02 port Cable6/0/5**	**Node is 3.8% degraded. Cacti Graphs**
18.4 MHz	**Upstream channel 0**	FEC problems - errored average of 0.6% (threshold 0.4%). Was in error 340 minutes (threshold 60 minutes).

Node NW060	**NWSTCMTK02 port Cable8/1/0**	**Node is 99.2% degraded. Cacti Graphs**
36.3 MHz	**Upstream channel 3**	FEC problems - errored average of 2.9% (threshold 0.4%). Was in error 1435 minutes (threshold 60 minutes).
15.9 MHz	**Upstream channel 4**	FEC problems - errored average of 9.3% (threshold 0.4%). Was in error 1410 minutes (threshold 60 minutes).
29.8 MHz	**Upstream channel 2**	FEC problems - errored average of 2.4% (threshold 0.4%). Was in error 1435 minutes (threshold 60 minutes).
23.3 MHz	**Upstream channel 1**	FEC problems - errored average of 2.7% (threshold 0.4%). Was in error 1435 minutes (threshold 60 minutes).

Node NW456	**NWSTCMTK02 port Cable6/1/2**	**Node is 17.4% degraded. Cacti Graphs**

Figure 12 – Example daily report output

Figure 12 is from the emailed report, showing a few different nodes with issues. One aspect that turned out to be very useful is this report, combined with Cacti's time-of-day logging on the graphs. A number of issues that had been plaguing different nodes were very time-of-day focused, with root causes ranging from irrigation sprinklers going off, to sunlight striking equipment at certain angles. The report highlighted the nodes to look at, and the charts allowed field techs to determine what time they needed to be at the node to start looking at the recurring problem.

One number that needed to be determined was "how bad is the node?", which is shown in Figure 12 in the "degraded" percentage. The formula for this was tweaked by hand to the current method:

> For each upstream, for each five minute polling interface, is the FEC percentage below 0.4%, is the traffic below the 80% threshold, and is the MER above 28.0 dB? If the answer is "no" to any question, consider the channel degraded during that interval. End by totaling up all the intervals that were degraded, divide by 288 for the number of intervals, and divide by the total number of upstreams to get the percentage.

As time has gone on and the issues have been corrected, these values and limits have been made less tolerant – at the start of the process, the FEC cutoff was up at 1.0%, for example.

This information is then summarized in a weekly report, sent to managers, directors, and VPs, containing a very high level summary of the issues for the week. See Figure 13.

STC	Nodes in STC	Number of nodes ever in error	Average number of days in error	Nodes in error 7+ days	Nodes continuously in error over the last 7 days	Nodes continuously in error over the last 14 days
ALIA	127	65 (51.2%)	3.3	23 (18.1%)	5: NQU81 NLV96 NNV58 NQW07 NQV84	2: NQU81 NNV58
CEHN	30	11 (36.7%)	2.3	3 (10.0%)	0:	0:
CENT	275	190 (69.1%)	6.3	91 (33.1%)	23: RON54 RQQ34 RQN59 RPQ53 ROT89 ROP02 RNN70 RSQ33 RSP09 RPQ19 RPQ18 RRQ47 RRN43 RSO83 RQN92 RPQ98 RRN30 RSN13 RPN26 RON19 ROL32 RSQ84 RQQ70	10: RON54 RQQ34 RQN59 RPQ53 ROT89 RSQ33 RPQ19 RRN43 RPQ98 RON19

Figure 13 – Example weekly report output

The weekly report is more focused on identifying nodes that have long-standing, long-running issues, to allow additional time and resources to be dedicated to fixing them. Other charts are included with the report (see Figures 14 and 15).

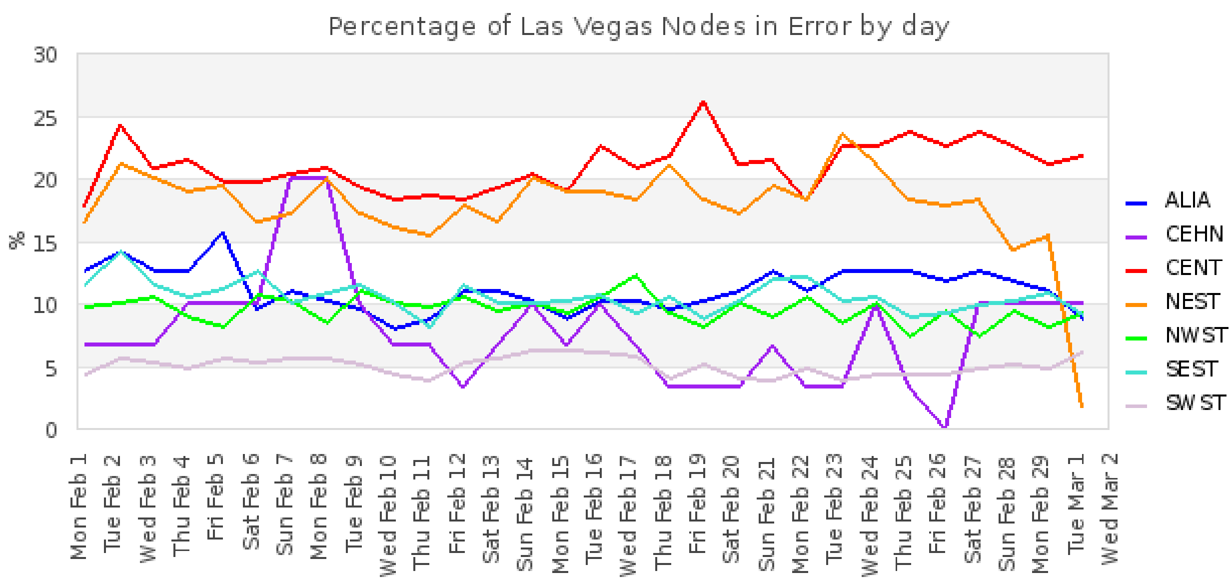

Figure 14 – Nodes in error summary graph

The chart in Figure 14 shows, by hub site, the percentage of problems existing. In this case, CENT and NEST have some of the oldest physical plant and highest percentage of aerial cabling, while SWST and NWST have a large percentage of new build-out.

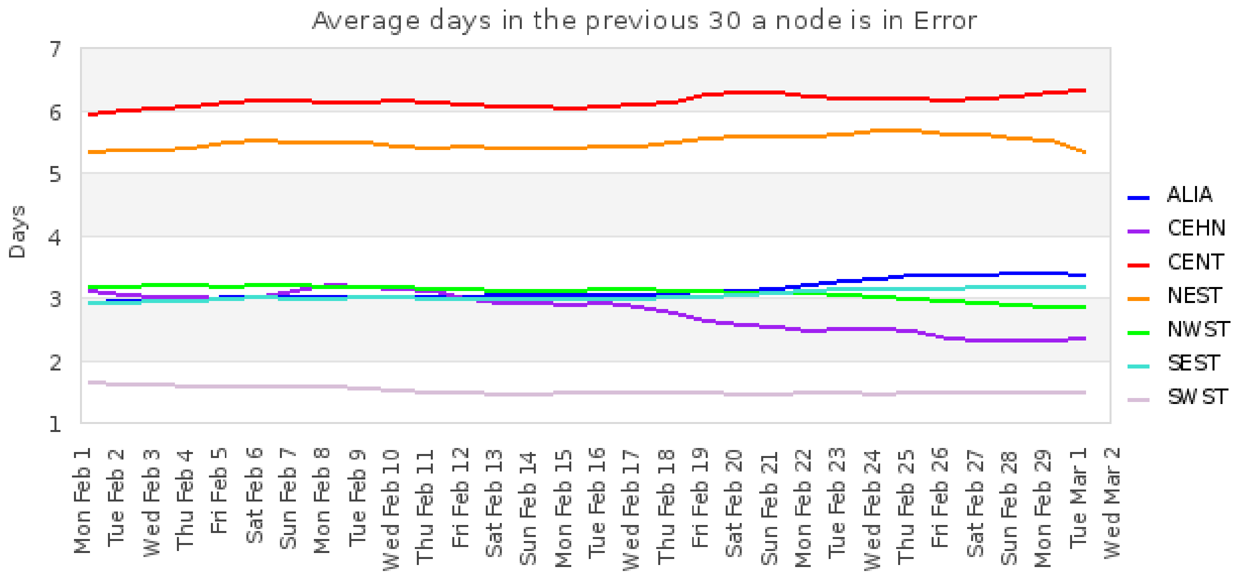

Figure 15 – Nodes in error duration summary graph

Figure 15 is an example of another chart that was added due to the curiosity of a director, and while the values in the chart are not very exciting, a longer story is involved. When the chart was first created, it highlighted the facilities that had a large number of nodes that were remaining in error. Management took this information and formed dedicated teams that, armed with the information of exactly which nodes to look at, when, and on what QAM carriers, quickly started eliminating long-running issues.

Finally, the frequency and automated nature of the reporting provides daily feedback to managers as to how things are going. Multiple times a dedicated effort has been implemented in a part of the market specifically to remove nodes from this list, with a resulting substantial reduction in call-in rates, customer truck rolls, and an improvement in overall customer satisfaction with the service in those areas. The data has also been used to help with technician training – those techs assigned to clean up nodes and those nodes still showing up on the report received additional training and assistance, which helped increase the efficiency of the workforce overall. While all of this data is available in other places, it is the combination of it, along with the reporting, that makes this a heavily used system.

Conclusion

This system was created by allowing the engineers with the data to be combined with ones with scripting capability, and conversations directly with the people needing the information. This alignment allowed for a quick deployment of technology that has now been operating for several years, and remains useful due to the ease of seeing the data, the focused and targeted analysis of the information that allows action to be easily taken, and the almost immediate feedback on results. This system is designed to not drown those using it in the depth of information available from other systems – it is designed to highlight exactly what information is needed to those who are in a position to do something about it.

Abbreviations

DOCSIS	Data-Over-Cable Service Interface Specifications
FEC	forward error correction
HFC	hybrid fiber-coax
ISBE	International Society of Broadband Experts
MER	modulation error ratio
MTA	multimedia terminal adapter
SCTE	Society of Cable Telecommunications Engineers
SNMP	simple network management protocol
STB	set top box