A Lesson in Eating Your Own Dog Food

April 18th 2011, Internet2 Spring Member Meeting – Network Tools Tutorial
Jason Zurawski – Internet2
The story you are about to hear is true. The names *have not* been changed to expose the guilty 😊

The following narrative exposes some fundamental flaws in the process of monitoring a network and learning to trust in available software tools and processes.

Recommendations for changing this behavior should be discussed.

Viewer discretion is advised.
Original Report – The Network is Broken!

- Feb 10th 2011 – Original report from Vanderbilt University (US CMS Heavy ION Tier2 Facility, Nashville TN) noting problems to Port d'Informació Científica (PIC – Barcelona Spain)
  - Concern #1 – Report was to me directly, does not scale

- Observation from users:
  - We are having trouble (slow transfers) with transfers from the CMS T1 sites in Spain (PIC). Here are traceroutes ... who can I talk to about this? Are we at least going along reasonable routes?

- Quick mental triage on my part:
  - Users are sharp, they have done this sort of thing before
  - They know the value of monitoring, and know when they are in over their head
  - Traceroutes are good, some real measurements would be better
  - Will require allocation of resources to address, coordinated by me now 😊
“I wish someone would develop a framework to make this easier”

- Yes, perfSONAR works well – *when it is deployed*.
- We still don’t have universal deployment, so the backchannel network of emails to “people you know” is still required

Coordination in domains, need to talk to people in each and allocate testers (if they don’t exist yet)

- PIC*
- CESCA
- RedIRIS
- GÉANT
- Internet2*
- SOX
- Vanderbilt*

* Started with these for simplicity
Resource Allocation & Instrumentation

• End Systems @ PIC and Vanderbilt
  – pS Performance Toolkit on a spare server
  – Racked next to the data movement tools
  – Benefits:
    • The similar OS and performance settings on each end “levels the playing field”
    • All tools are now available, if we want to run an NDT we can, if we need regular BWCTL, we have it.
      – Cost to me and remote hands = < 1hr of installation/configuration

• Internet2
  – Regular BWCTL, OWAMP testing in place.
  – Interface Utilization and Errors available for all links
  – Web100 enabled services for NDT and NPAD
  – Cost to me = free
Long Path Testing

- **Sanity Checking**
  - Way to verify that the problem is not with the application (e.g. in this case it was a CMS product – PhEDEx)
  - Establish some baselines with all tools
  - Try different times of day to rule out congestive issues

- **Initial Results:**
  - NDT confirms poor performance (100s of Kbps) in the PIC to Vanderbilt direction. Other way is not bad (500Mbps).
  - BWCTL is similar, even when using parallel streams
  - Traceroutes show a route asymmetry (although in practice this alone is not enough to cause this large of a problem...):
    - PIC->CESCA->RedIRIS->GEANT->Internet2->SOX->Vanderbilt
    - Vanderbilt->SOX->NLR->GEANT->RedIRIS->CESCA->PIC
  - Time of day had no serious effect, this appears to be a real problem.
**Parallel Effort**

- PhEDEx has its own notion of the performance it gets, and the Vanderbilt and PIC operators know more about this than I ever will
  - Continued to ‘tweak’ application settings to see if they could do better than my tools
  - Still didn’t get very far

- Sidebar – PhEDEx is pretty intelligent
  - Will alter its behavior due to historical results
  - Takes advantage of fast paths
  - Does this over time, not instant
Structured Debugging Methodology

- Divide and Conquer
  - Bisect the path and test the segments individually
  - Rule out paths that are doing well, subdivide those that aren’t again and again

- Use of one tool a time
  - Collect as much as you can with each tool
  - Move to the next to gather different metrics

- Patience
  - It’s not hard, but it is time consuming

- If only we had a way to teach this to people...
  - Wait ... we do!
  - [http://www.internet2.edu/workshops/npw/](http://www.internet2.edu/workshops/npw/)
Real Debugging – Timeline and Results

- Initial Report: Feb 10th 2011
- Structured Debugging: Feb 23rd – 24th 2011

- Methodology
  - GÉANT Circuit from Frankfurt terminates at Internet2 Washington DC. Use test points here.
  - Vanderbilt connects through SOX, which connects to Internet2 in Atlanta GA. Use test points here too.
  - 2 10G backbone links separate Atlanta and Washington.
Real Debugging – Results (Traceroutes)

- Between PIC and Vanderbilt were asymmetric
  - PIC->CESCA->RedIRIS->GEANT->Internet2->SOX->Vanderbilt
  - Vanderbilt->SOX->NLR->GEANT->RedIRIS->CESCA->PIC

- Focus on the US connectivity:
  - Between Vanderbilt and 2 Internet2 hosts, no asymmetry was observed
  - Path:
    - Vanderbilt->SOX->Internet2 (ATLA)->Internet2 (WASH)
Real Debugging – Results (Utilization)

- In the Internet2 case, utilization and errors are available.
- There are two backbone links between ATLA and WASH
  - 10G CPS Link – ruled this out of the process
  - 10G R&E Link
NDT is not run “regularly”, so our use will strictly be diagnostic.

- Vanderbilt (client) -> PIC (server)
  - running 10s outbound test (client to server) . . . . 522.24 Mb/s
  - running 10s inbound test (server to client) . . . . 169.89 kb/s

- Vanderbilt (client) -> WASH (server)
  - running 10s outbound test (client to server) . . . . 922.47 Mb/s
  - running 10s inbound test (server to client) . . . . 1.35 Mb/s

- Vanderbilt (client) -> ATLA (server)
  - running 10s outbound test (client to server) . . . . 935.98 Mb/s
  - running 10s inbound test (server to client) . . . . 933.82 Mb/s
Real Debugging – Results (NDT, cont.)

• We now have a minor result
  – Performance on a shorter path to from Vanderbilt to ATLA seems expected.
  – Can we use this to our advantage?
• Internet2 Atlanta (client) -> Internet2 Washington (server)
  – running 10s outbound test (client to server) . . . . 978.44 Mb/s
  – running 10s inbound test (server to client) . . . . 251.95 kb/s
• Very promising result ... but we aren’t done!
  – Can’t declare victory with just this
  – Use other tools as much as we can
  – See if we can confirm that this segment is a problem
• Reality Check
  – Concern #2 – This could potentially be a backbone link ... how could something that is monitored 24/7/365 be in that poor of a state ...
Real Debugging – Side Bar

- Related information is a good thing. There is a trouble ticket system that alerts to changes in the network:

Subject: Internet2 Layer 1 Various Circuits (WASH) Maintenance

Affected:
- I2-ATLA-WASH-I2-05132
- I2-ATLA-WASH-I2-05247
- I2-CHIC-WASH-I2-05246
- I2-CHIC-WASH-I2-05382
- I2-CLEV-WASH-I2-08082
- I2-NEW32A0A-WASH-I2-04641
- I2-NEW32A0A-WASH-I2G-E85243
- I2-PHI-WASH-I2-08792
- I2-PITT-WASH-I2-08290
- I2-PITL-WASH-I2-05791

Scheduled Start Time: Wednesday, February 23, 2011, 4:00 AM (0800 UTC)
Scheduled End Time: Wednesday, February 23, 2011, 8:00 AM (0800 UTC)
Description: The above listed circuits via Washington, D.C. will be unavailable to the community while Internet2 Engineers swap the MX 960 to a T1K00. An extended outage is expected; however, the entire window is reserved.

Ticket No.: 1493-100
Timestamp: Tue Feb 22 04:53:13 2011 UTC

When replying, type your text above this line.
Please submit problems, requests, and questions at:
http://noc.net.internet2.edu/i2network/support/report-a-problem.html

Thank You,
The Internet2 Network Operations Center
Indiana University
noc@net.internet2.edu, 317-278-6622

Internet2 NOC Home Page: http://noc.net.internet2.edu/
Internet2 NOC Operations Calendars (RSS and iCAL):
http://noc.net.internet2.edu/i2network/support/operations-calendar.html
Internet2 NOC Notification listserv Archive:
https://listserv.indiana.edu/archives/internet2-ops-1.html
Internet2 NOC Weekly Reports:
http://noc.net.internet2.edu/i2network/support/weekly-reports.html
Regular monitoring is your friend ... WHEN YOU USE IT PROPERLY 😊

- Internet2 has lots of fancy GUIs that expose the BWCTL data, these should be viewed every now and then
- We even have plugins for NAGIOS developed by perfSONAR-PS to alarm when performance dips below expectations
- We did neither of these properly 😊

### BWCTL - Internet2 Network IPv4 TCP Throughput

<table>
<thead>
<tr>
<th>bwctl/ipert</th>
<th>Atlanta</th>
<th>Chicago</th>
<th>Houston</th>
<th>KansasCity</th>
<th>LosAngeles</th>
<th>NewYorkCity</th>
<th>SaltLakeCity</th>
<th>Seattle</th>
<th>Washington</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44.93</td>
</tr>
</tbody>
</table>
Real Debugging – Results (BWCTL)

- Digging Deeper on WASH:

---

<table>
<thead>
<tr>
<th>Source: Washington</th>
<th>Destination: Atlanta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>988.26 Mbps</td>
</tr>
<tr>
<td>Average</td>
<td>862.92 Mbps</td>
</tr>
<tr>
<td>Last</td>
<td>47.17 Mbps</td>
</tr>
<tr>
<td>Destination</td>
<td>Maximum</td>
</tr>
<tr>
<td></td>
<td>988.43 Mbps</td>
</tr>
<tr>
<td></td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>988.25 Mbps</td>
</tr>
<tr>
<td></td>
<td>Last</td>
</tr>
<tr>
<td></td>
<td>988.43 Mbps</td>
</tr>
</tbody>
</table>

---

Graph showing bandwidth graph for perfSONAR-PS perfAdmin Bandwidth Graph with source and destination data.
Real Debugging – Results (BWCTL)

- Remember that trouble ticket ...
Real Debugging – Results (BWCTL)

• Concern #3 – This poor performance was detected by the tools, but went undetected by the humans in operations. Need to change the culture of operations - learn to listen to (and trust!) the tools
Now we have several results

- NDT diagnostics show poor results
  - PIC->Vanderbilt
  - WASH->Vanderbilt
  - WASH->ATLA
- NDT diagnostics show good results
  - ATLA->Vanderbilt
- BWCTL regular monitoring shows poor results
  - ATLA to WASH
  - ATLA to NEWY (which goes over the WASH path), we can ignore further debugging for here for now
- BWCTL regular monitoring shows good results
  - Everywhere else
- Don’t call it a day yet! One more tool too look at.
Real Debugging – Results (OWAMP)

- Much like BWCTL, we keep this going all the time. Also like BWCTL, we don’t have alarms to tell us things are bad 😊
Interpreting the graph shows a pretty constant stream of loss (WASH -> ATLA). Note this is a “soft failure”, not loss of connectivity.
Problem Location and Reporting

- At this stage we have our evidence from all of the tools.
- Time to escalate – this is why we have 24/7/365 NOCs after all
- Problem reported Feb 24th 2011 @ 4pm EST
  - Evidence from tests above provided, lots of detail!
- First response from operations:

  > Jason,
  >
  > I'm not seeing any degradation over that backbone link. Could you provide me with a trace? Perhaps there is loss occurring elsewhere along the path?
  >
  > Thanks,
  >
  > Greg
Denial?

• This first response should make anyone upset, especially after providing evidence from four (4!) tools
  – Concern #4 – Listening to the tools is one problem, **TRUSTING** them is another that needs to be overcome in the operations space

• To be fair ... operations may have a different set of tools they are working with:
  – Monitoring of the Interface counters is something most are taught to watch – we revealed on Slide 11 that there was no evidence of errors. Utilization looked “ok”
  – Can’t speak for the regular monitoring – these have been in place on the Internet2 observatory for around 6 years. Alarming is not in place at a minimum.
Ok, Not Complete Denial...

- Bringing in more eyes sometimes gets results, especially when they have looked at the evidence and can agree something doesn’t smell right...

---

Entered on 02/24/2011 at 22:32:07 UTC (GMT+0000) by Tom Knoeller:

We are seeing errors increasing on the ATLA side. But light levels look good on both sides. Probably need to do some emergency work to throw some loops in the circuit to see where the problem is. My guess is going to be the XENPAK PIC on the WASH router as that is what changed 2 days ago, but testing will confirm that.

Traffic on the link is light, so I think we can turn off ISIS to divert the traffic without to much pain.

Service Desk: Lets get this into the hands of the oncall to work tonight.

Thanks,
-Tom
Testing Hypothesis

- When operations tweaks things, the tools know:
Testing Hypothesis

- Interpretation:

![Graph showing Interface Bounce](image-url)
Testing Hypothesis

- Explanation from the trouble ticket:

> Traffic on the link is light, so I think we can turn off
> ISIS to divert the traffic without to much pain.

And for those playing the home game, we tried to turn off the link, but
I did not think about the offered load being higher with no packet loss,
so it overloaded the other backbone link. At this point, the interface
is turned on and running in a degraded state until a emergency FSR can
be done to move to a new PIC.

-Tom
Testing Hypothesis

- Next step:

- Maintenance was scheduled for Feb 24th 2011 @ 6:30PM EDT
  - If you are keeping track, this is only 2.5 hours since the ticket was opened

Entered on 02/24/2011 at 23:47:08 UTC (GMT+0000) by Hans Addleman:
Tom suspected and I agreed that this might be the XENPAK optic failing in the T1600. I had a tech run a new fiber over to a new port (1/1/3) and the errors are still being observed on the ATLA side.

Next course is going to be terminal looping the circuit and doing some testing to see if this is a layer 1 issue perhaps.

Hans Addleman
IU Global NOC Engineer
addlema@grnoc.iu.edu
Solution In Place ... Will It Hold?

- Not longer after swapping to a different interface:

Entered on 02/25/2011 at 00:02:08 UTC (GMT+0000) by Hans Addleman:
Okay.. it just took a minute for the counters to settle down.
The swap of interfaces fixed the problem! Traffic on that link jumped up by almost 3gig and the link looks healthy again.
So we have a bad xenpak in WASH that we can worry about in the morning. Tom is going to work with Ross to start sending out spares to the sites.
Thanks to Tom for all the initial leg work on this.. Made my part of it this evening very easy.
Hans Addleman
IU Global NOC Engineer
adlema@grnoc.iu.edu

- And what do the tools say ...
Solution In Place ... Will It Hold?

- OWAMP is sensitive, so let's go back to it:
Solution In Place ... Will It Hold?

- Interpreting:

![Graph showing data points and labeled areas: Loss, No More Loss, Interface Switch]
Solution In Place ... Will It Hold?

- What about BWCTL?
Solution In Place ... Will It Hold?

- Interpreting:
Solution In Place ... Will It Hold?

- Lastly, how about network utilization. In theory this should have limited all traffic...
Solution In Place ... Will It Hold?

- And it did ...

Traffic Improves
Re-testing, Notification of Customer

- NDT is good for a one off, lets verify the paths again
- Vanderbilt (client) -> WASH (server)
  - running 10s outbound test (client to server) . . . . . . 923.47 Mb/s
  - running 10s inbound test (server to client) . . . . . . 914.02 Mb/s
- Vanderbilt (client) -> PIC (server)
  - running 10s outbound test (client to server) . . . . . . 524.05 Mb/s
  - running 10s inbound test (server to client) . . . . . . 550.64 Mb/s
- Not “perfect”, but closer
  - Client was asked to verify CMS applications
  - Debugging shouldn’t stop, there are more parts of the path to explore.
Lessons Learned/Next Steps

The 4 concerns:
- Reporting problems – who should “own” them? The campus? The RON? The Backbone? An organization, e.g. a PERT, organized for this type of thing?
- State of a major link – how could this error have gone unnoticed by the watchers. It was good the user spoke up…
- Using the tools in a operations environment. These are not just toys – the goal is to convert R&D activities to real applications
- Trusting the tools, if they are in fact untrustworthy, we should be working to make them better.
Final Thoughts

• Fast in comparison to other problems:
  – Initial Report = Feb 11th, one problem corrected Feb 24th (10 Business Days)
• Streamlining:
  – Tool Ubiquity
  – Built into NOC process
• Value of tools shown
  – NOC Staff
  – Customers (on two continents!)
• Where do we go from here?

Invisible High Five
A Lesson in Eating Your Own Dogfood

April 18th 2011, Internet2 Spring Member Meeting – Network Tools Tutorial
Jason Zurawski – Internet2

For more information, visit http://www.internet2.edu/performance