An Analysis of Network Configuration Artifacts

LISA '09, November 5, 2009

David Plonka & Andres Jaan Tack
{plonka,tack}@cs.wisc.edu
Motivation and Goals

- Like software quality, network reliability is evolving:
  - Expectation of high availability, increasing reliance
  - Increasing numbers of skilled practitioners
  - Increasing level of automation
Motivation and Goals

• Like software quality, network reliability is evolving:
  – Expectation of high availability, increasing reliance
  – Increasing numbers of skilled practitioners
  – Increasing level of automation

• However, the management of networks and the Internet has not received similar attention to the development of software.
Motivation and Goals

- Like software quality, network reliability is evolving:
  - Expectation of high availability, increasing reliance
  - Increasing numbers of skilled practitioners
  - Increasing level of automation
- However, the management of networks and the Internet has not received similar attention to the development of software.
- We propose an analogy-based analysis, and that these elements are akin to each other:
  - Networks : Software Systems
  - Network Engineering : Software Engineering
  - Network Operators : Programmers
Campus Network
Network Artifacts

• *artifact* - an object created by humans, especially one remaining from a particular period

• Network Performance Measurements

• Network Management Systems' Topology

• Trouble Tickets

• **Network Device Configurations**
  - Routers, switches, firewalls
  - Network practitioners use Source Code Management (SCM) of device configurations for:
    - Configuration backups
    - Communicating changes
Network Configuration Repositories
## Networks Studied

<table>
<thead>
<tr>
<th>Network</th>
<th>Period in Years</th>
<th>Operators (super-users)</th>
<th>Devices / Configuration Files</th>
<th>Revisions</th>
<th>Lines of Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campus</td>
<td>5+</td>
<td>343 (64)</td>
<td>3,839</td>
<td>128,394</td>
<td>2,898,362</td>
</tr>
<tr>
<td>Service Provider</td>
<td>10+</td>
<td>31 (31)</td>
<td>519</td>
<td>41,787</td>
<td>163,882</td>
</tr>
</tbody>
</table>
Mining SCM Repositories - Why?

- While successful in the PL community, this hasn't been leveraged in the context of network configuration and management.
- To visualize and elucidate network operation with the goal of understanding and improving the practice.
Mining SCM Repositories - How?

- Convert existing custom network version control system repositories to common CVS repositories.
- Use existing tools from the Programming Language (PL) and open source developer communities, e.g.:
  - StatCVS-XML
  - cvs2cl (CVS to ChangeLog)
- Perform additional static file analyses, e.g.:
  - Syntax-aware statistics (i.e. config stanzas)
  - Revision lifetimes
version 12.2
no service pad
service timestamps debug datetime localtime
service timestamps log datetime localtime
service password-encryption
!
hostname s-bldg-5-2-access
!
spanning-tree mode rapid-pvst
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
!
interface FastEthernet1/0/1
  description sample 100Mbps ethernet interface
  switchport access vlan 42
  switchport mode access
  ip access-group nodhcpserver in
  snmp trap mac-notification change added
  snmp trap mac-notification change removed
  no snmp trap link-status
  no mdix auto
  spanning-tree portfast
  spanning-tree bpduguard enable
  spanning-tree guard root
ip access-list extended nodhcpserver
remark Id: ndhcp.acl,v 1.2 2005-05-20 11:26:03 ashley Exp
deny udp any eq bootps any
permit ip any any
!
access-list 5 permit 192.2.0.1
access-list 5 remark Allow foo, bar, and baz servers
access-list 5 permit 192.2.0.10
access-list 5 permit 192.2.0.11
!
...
Campus File / Device Count

Campus Network
File Count

Files
0
250
500
750
1000
1250
1500
1750
2000
2250
2500
2750
3000
3250
3500
3750
4000

Date
Jan-1999
May-1999
Sep-1999
Jan-2000
May-2000
Sep-2000
Jan-2001
May-2001
Sep-2001
Jan-2002
May-2002
Sep-2002
Jan-2003
May-2003
Sep-2003
Jan-2004
May-2004
Sep-2004
Jan-2005
May-2005
Sep-2005
Jan-2006
May-2006
Sep-2006
Jan-2007
May-2007
Sep-2007
Jan-2008
May-2008
Sep-2008
Jan-2009
Campus LOC by Topology

Campus Network
Lines Of Code (per module)

campus/access/ = 1,912,430

campus/access/wireless/ = 601,836

campus/firewall/ = 120,147

campus/mgmt/ = 117,756

campus/dist/ = 98,921

campus/core/ = 47,272
Campus Network
Lines Of Code (per author)

Campus Size Per Author

alexander  ann  annie  antonio  betty  bradley  cathryn  christina  cindy  daniel  danny  debbie  deborah  don  edwin
elizabeth  frederick  glenn  grace  jason  jim  jimmy  joe  kenny  juanita  judith  kathryn  kelly  kimberly  leonard
michael  michelle  mike  monica  paula  ray  raymond  renee  rhonda  samuel  shannon  steve  steven  tiffany  tom
tracy  travis  troy  victor  virginia  wayne  ann  barry  catherine  dana  danielle  diane  howard  janice  jay  jeffery
jeffrey  john  jonathan  justin  lauren  margaret  maria  matthew  mildred  phyllis  rebecca  sharon  stanley  aaron
amy  brian  ellen  joseph  josephine  nicole  pamela  patrick  randy  russell  ruth  timothy  valerie  vincent  craig
douglas  gail  gloria  kathleen  linda  michelle  nicholas  samantha  therma  victoria  wendy  amber  brenda  carmen
clarence  david  edna  jeremy  joan  jian  lisa  thomas  alice  doris  jackie  marcus  mark  norma  yvonne  amanda
anne  jamie  lawrence  nancy  robert  sylvia  theodore  julia  ana  carolyn  barbara  jacqueline  bobby  jack  laura
rosa  shawn  heather  michael  may  sandra  sheila  vivian  donna  carl  charles  geraldine  francis  lynn  william
albert  carol  sherry  roy  stacy  pauline  wanda  christopher  eleanor  henry  christine  anita  judy  andrea  helen
billy  randall  philip  crystal  gladys  darlene  luis  jacob  jean  alan  miguel  charlotte  clara  donald  george
jessica  martin  herbert  francis  jeff  james  megan  veronica  tina  jill  harold  jane  fred  jerry  hazel  chris
susan  scott  tony  eugene  leroy  joshua  erin  todd  richard  rita  anthony  ryan  marvin  kenneth  carrie  eva
norman  audrey  eric  angela  ethel  johnny  walker  ronald  regina  keith  bruce  andrew  edward  shirley  lois
gregory  melvin  florence  anna  emma  gerald  april  brandon  ruby  edith  nathan  peggy  sara  ashley  kevin
Campus Size Per Group

Campus Network by Device Type
Lines Of Code (per author)

- net = 2,418,758
- contract = 348,207
- noc = 100,099
- field = 57,582
- authorized-agents = 1,821
- security = 103

Legend:
- security
- authorized-agents
- field
- noc
- contract
- net
Campus Commits by Hour
## Common Commit Comments

<table>
<thead>
<tr>
<th>Comment</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial revision</td>
<td>1487 (2.7%)</td>
</tr>
<tr>
<td>test</td>
<td>812 (1.5%)</td>
</tr>
<tr>
<td>asdf</td>
<td>593 (1.1%)</td>
</tr>
<tr>
<td>'newer bulk checkin'</td>
<td>411 (0.7%)</td>
</tr>
<tr>
<td>change vlan</td>
<td>316 (0.6%)</td>
</tr>
</tbody>
</table>
## An Anomaly

<table>
<thead>
<tr>
<th>Author</th>
<th>Revisions</th>
<th>Lines of Code</th>
<th>Added Lines of Code</th>
<th>Lines of Code per Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>net</td>
<td>63468 (47.2%)</td>
<td>2418758 (82.9%)</td>
<td>3313853 (74.1%)</td>
<td>38.11</td>
</tr>
<tr>
<td>authorized-agents</td>
<td>38625 (28.8%)</td>
<td>1821 (0.1%)</td>
<td>208956 (4.7%)</td>
<td>0.05</td>
</tr>
<tr>
<td>system</td>
<td>11218 (8.4%)</td>
<td>-8795 (-0.3%)</td>
<td>125618 (2.8%)</td>
<td>-0.78</td>
</tr>
<tr>
<td>noc</td>
<td>10715 (8.0%)</td>
<td>100099 (3.4%)</td>
<td>303481 (6.8%)</td>
<td>9.34</td>
</tr>
<tr>
<td>field</td>
<td>6122 (4.6%)</td>
<td>57582 (2.0%)</td>
<td>152498 (3.4%)</td>
<td>9.41</td>
</tr>
<tr>
<td>contract</td>
<td>3959 (2.9%)</td>
<td>348207 (11.9%)</td>
<td>368518 (8.2%)</td>
<td>87.95</td>
</tr>
<tr>
<td>security</td>
<td>230 (0.2%)</td>
<td>103 (0.0%)</td>
<td>1898 (0.0%)</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Code Push Results

Made directory: /home/net/cms/codepusher/af4f5bf32242246a603eb619ce60cc66

Created working set file.
Push is running in process 5604.

Creating command files...
Created command file for s-1301ua-214-1-access.
Created command file for s-agentl-120a-1-access.
Created command file for s-cscc-b295-2-access.

Done creating command files.

Making upgrade.make... done!

Running upgrade.make
s-1301ua-214-1-access.log... starting.
s-agentl-120a-1-access.log... starting.
s-cscc-b295-2-access.log... starting.
s-cscc-b295-2-access.log... writing.
s-agentl-120a-1-access.log... writing.
s-1301ua-214-1-access.log... completed. Time elapsed: 4 wallclock secs
s-1301ua-214-1-access.log... completed. Time elapsed: 4 wallclock secs
s-cscc-b295-2-access.log... completed. Time elapsed: 4 wallclock secs

Done pushing code!

COMMANDS PUSHED:

show boot.

Beginning error scan...
Scanning s-1301ua-214-1-access.log for errors... No errors found.
Scanning s-agentl-120a-1-access.log for errors... No errors found.
Scanning s-cscc-b295-2-access.log for errors... No errors found.

Error scan complete

Code push completed.
Your log files are located in /home/net/cms/codepusher/af4f5bf32242246a603eb619ce60cc66
You should check them manually for errors.

Check in your config files, if necessary!!

AANTS Home Page | EdgeConf | UW Home

If you have feedback or questions, please contact the aants administration team at:

sants-admin@net.dotl.wisc.edu

Copyright © 2004 The Board of Regents of the University of Wisconsin System
Evaluating Practitioner Effort

- Measurements of practitioner effort
  - How often are “fixes” introduced?
  - How often do configurations change?
  - “Bad Days” (are Friday checkins more buggy?)

- Look toward improvements:
  - Syntax-aware revision analysis (stanzas)
  - How do we direct tool development?
Revision Lifetimes

• How long does a revision last before it is next modified?
  – Suggests the modus operandi of practitioners
  – Suggests the value or the staying power of a revision
  – Might also suggest some measure of network volatility
Campus Revision Lifetimes (<3.5 days)
Campus Revision Lifetimes (<10 min)
% Short-Lived Revisions by Day

- Sun
- Mon
- Tue
- Wed
- Thu
- Fri
- Sat

Service Provider
Campus
% Short-Lived Revisions by Day

(Campus Drill-Down)
Campus Average File Size

Campus Network
Average Filesize

Date
LOC / File
250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450 1500
Service Provider Average File Size

Service Provider Network
Average Filesize

LOC / File

Date
## Campus Revisions by Stanza Type

<table>
<thead>
<tr>
<th>Stanza Type</th>
<th>Total Revision Count</th>
<th>Revisions per Instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>interface</td>
<td>471,238</td>
<td>4</td>
</tr>
<tr>
<td>vlan</td>
<td>25,591</td>
<td>1</td>
</tr>
<tr>
<td>global</td>
<td>12,534</td>
<td>4</td>
</tr>
<tr>
<td>logging</td>
<td>12,390</td>
<td>9</td>
</tr>
<tr>
<td>ip</td>
<td>12,006</td>
<td>1</td>
</tr>
<tr>
<td>bridge</td>
<td>4,353</td>
<td>1</td>
</tr>
<tr>
<td>line</td>
<td>3,936</td>
<td>1</td>
</tr>
<tr>
<td>banner</td>
<td>3,810</td>
<td>1</td>
</tr>
<tr>
<td>dot11</td>
<td>3,324</td>
<td>1</td>
</tr>
<tr>
<td>control-plane</td>
<td>3,013</td>
<td>1</td>
</tr>
</tbody>
</table>
Some Conclusions

- With varying device types, LOC is an erratic metric for the stanza-based, declarative network configuration language, (such as Cisco IOS)
- Analysis of network configurations exposes pertinent network management details including:
  - Group behaviors
  - Outstanding practitioners
  - Change times
  - High level of user compliance, but some curiosities
  - Tool-based efficiencies both expected and invented
Contributions

• An initial application of software development analysis tools to network operations based on existing, freely-available tools

• Beginnings of a network operations-specific measurement of practitioner effort to guide tool development, such as SCM and IDE-like tools for network operators

• In our case studies, this analogy-based analysis approach shows promise based on feedback by expert interviews.
Discussion and Future Work

- As in software, can we identify and investigate code decay, refactorings, and code clones?
- Leverage other artifacts to measure practitioner compliance and network service reliability and performance.
- Develop a complexity metric based on stanzas and inter-stanza references. (see Benson, et al., NSDI 2009)
An Analysis of Network Configuration Artifacts

LISA '09, November 5, 2009

David Plonka & Andres Jaan Tack
{plonka,tack}@cs.wisc.edu