Curbing Abusive Behavior on Science Gateways

Pascal Meunier and Michael McLennan, Purdue University

NSF Cyber Security Summit 2014

60+ Hubs

hubzero

open source™ Software Platform for Collaboration

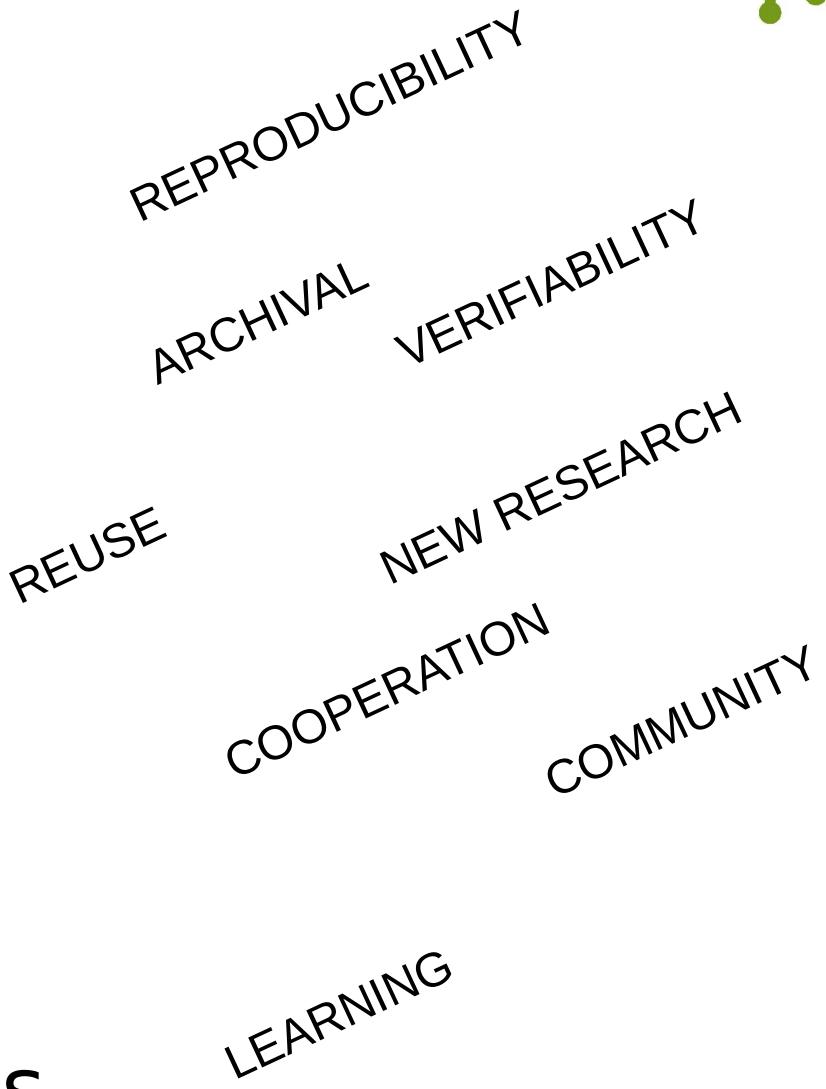
~1,500,000

visitors total

visitors	users	
689,743	330,251	nanoHUB.org
343,350	112,862	nees.org
64,131	32,763	pharmaHUB.org
59,517	4,669	HABRIcentral.org
56,355	14,646	vhub.org
47,967	23,088	GlobalHUB.org
46,710	12,643	cceHUB.org
44,723	5,372	PURR
41,689	5,396	iemhub.org
40,289	8,207	StemEdHub.org
39,188	6,362	ciHUB.org
39,134	7,933	molecularHUB.org

Hubs Provide

- Tools
  - Calculations
  - Simulations
  - Vizualization
  - Code
- Projects, groups
- Videos and slides
- References
- Classes and Tutorials



Problem Statement

- Want:
  - Open and easy to use
    - First use
  - Features
    - Useful and engaging
  - Media uploads and downloads
  - Low maintenance costs

- Abuse
  - Ads, Spam
  - Email
  - Links
  - Blog spam
  - HTML, images
  - Malware
  - Recon and Exploits
  - Storage of unrelated things
- Spend a lot of time cleaning, deactivating accounts and blocking

Problem Statement

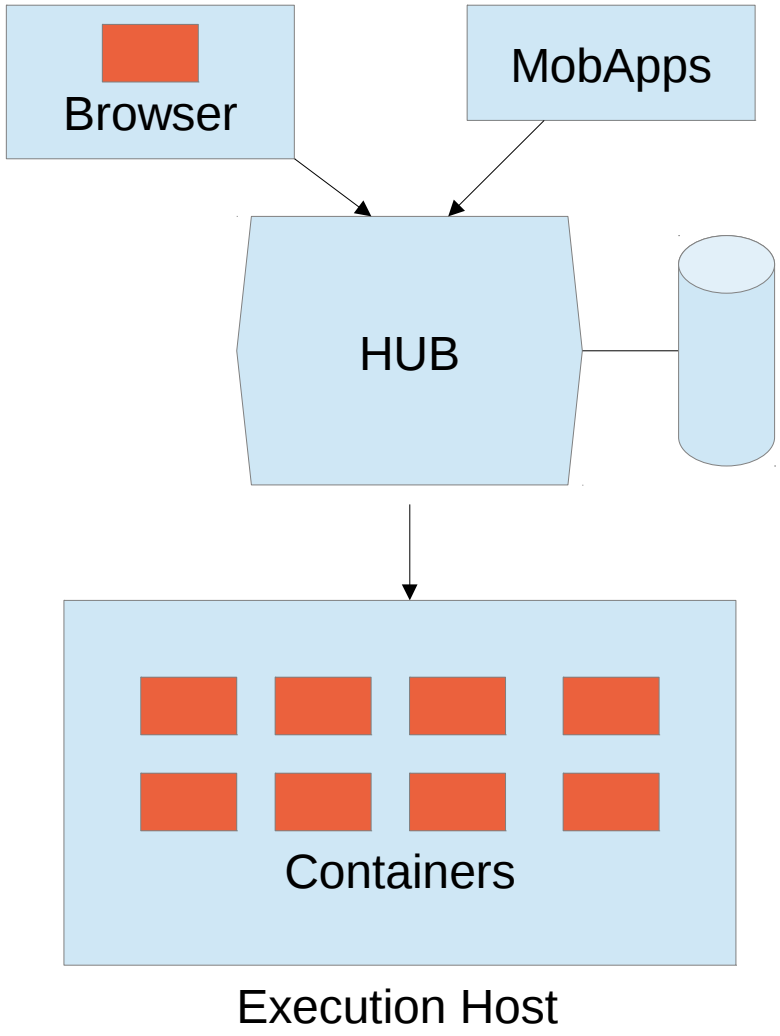
- We authenticate but...
- Who are you really?
- We can't identify you solidly
- Email addresses are cheap
- Accounts have little value so abusers create them at will
- We can't complain about you
- Little accountability

What Didn't Work

- Web Application Firewall
  - Identify normal activity and let it through
  - Identify attacks and block them
  - Needs to model the web application
    - Complex application means a complex, customized WAF

What Didn't Work

- Modsecurity
  - HUBs are complex
    - Many false positives
  - Custom or uncommon functionality
    - VNC proxy
    - API
    - SVN
  - Many extremely complex regular expressions to manage
    - Introduce bugs
    - Time consuming
    - Truncate/simplify/lose



Example (one regex)


```
SecRule REQUEST_COOKIES|REQUEST_COOKIES_NAMES|REQUEST_FILENAME|ARGS_NAMES|ARGS|XML:/* "(?i:(?:(?:s(?:t(?:d(?:dev(_pop|_samp)?)?|r(?:_
to_date|cmp))|u(?:b(?:str(?:ing(_index)?)?|(?:dat|tim)e)|m)|e(?:c(?:_to_time|ond)|ssion_user)|ys(?:tem_user|date)|ha(1|2)?|oundex|che
ma|ig?n|pace|qrt)|i(?:s(null|_(free_lock|ipv4_compat|ipv4_mapped|ipv4|ipv6|not_null|not|null|used_lock))?|n(?:et6?_(aton|ntoa)|s(?:er
t|tr)|terval)?|f(null)?)|u(?:n(?:compress(?:ed_length)?|ix_timestamp|hex)|tc_(date|time|timestamp)|p(?:datexml|per)|uid(_short)?|case
|ser)|l(?:o(?:ca(?:l(timestamp)?|te)|g(2|10)?|ad_file|wer)|ast(_day|_insert_id)?|e(?:(?:as|f)t|ngth)|case|trim|pad|n)|t(?:ime(stamp|s
tampadd|stampdiff|diff|_format|_to_sec)?|o_(base64|days|seconds|n?char)|r(?:uncate|im)|an)|m(?:a(?:ke(?:_set|date)|ster_pos_wait|x)|i
(?:(?:crosecon)?d|n(?:ute)?)|o(?:nth(name)?|d)|d5)|r(?:e(?:p(?:lace|eat)|lease_lock|verse)|o(?:w_count|und)|a(?:dians|nd)|ight|trim|p
ad)|f(?:i(?:eld(_in_set)?|nd_in_set)|rom_(base64|days|unixtime)|o(?:und_rows|rmat)|loor)|a(?:es_(?:de|en)crypt|s(?:cii(str)?|in)|dd(?
:dat|tim)e|(?:co|b)s|tan2?|vg)|p(?:o(?:sition|w(er)?)|eriod_(add|diff)|rocedure_analyse|assword|i)|b(?:i(?:t_(?:length|count|x?or|and
)|n(_to_num)?)|enchmark)|e(?:x(?:p(?:ort_set)?|tract(value)?)|nc(?:rypt|ode)|lt)|v(?:a(?:r(?:_(?:sam|po)p|iance)|lues)|ersion)|g(?:r(
?:oup_conca|eates)t|et_(format|lock))|o(?:(?:ld_passwo)?rd|ct(et_length)?)|we(?:ek(day|ofyear)?|ight_string)|n(?:o(?:t_in|w)|ame_cons
t|ullif)|(rawton?)?hex(toraw)?|qu(?:arter|ote)|(pg_)?sleep|year(week)?|d?count|xmltype|hour)\W*\(|\b(?:(?:s(?:elect\b(?:.{1,100}?\b(?
:(?:length|count|top)\b.{1,100}?\bfrom|from\b.{1,100}?\bwhere)|.*?\b(?:d(?:ump\b.*\bfrom|ata_type)|(?:to_(?:numbe|cha)|inst)r))|p_(?:
sqlexec|sp_replwritetovarbin|sp_help|addextendedproc|is_srvrolemember|prepare|sp_password|execute(?:sql)?|makewebtask|oacreate)|ql_(?
:longvarchar|variant))|xp_(?:reg(?:re(?:movemultistring|ad)|delete(?:value|key)|enum(?:value|key)s|addmultistring|write)|terminate|xp
_servicecontrol|xp_ntsec_enumdomains|xp_terminate_process|e(?:xecresultset|numdsn)|availablemedia|loginconfig|cmdshell|filelist|dirtr
ee|makecab|ntsec)|u(?:nion\b.{1,100}?\bselect|tl_(?:file|http))|d(?:b(?:a_users|ms_java)|elete\b\W*?\bfrom)|group\b.*\bby\b.{1,100}?\
bhaving|open(?:rowset|owa_util|query)|load\b\W*?\bdata\b.*\binfile|(?:n?varcha|tbcreato)r|autonomous_transaction)\b|i(?:n(?:to\b\W*?\
b(?:dump|out)file|sert\b\W*?\binto|ner\b\W*?\bjoin)\b|(?:f(?:\b\W*?\(\W*?\bbenchmark|null\b)|snull\b)\W*?\()|print\b\W*?\@\@|cast\b\W
*?\()|c(?:(?:ur(?:rent_(?:time(?:stamp)?|date|user)|(?:dat|tim)e)|h(?:ar(?:(?:acter)?_length|set)?|r)|iel(?:ing)?|ast|r32)\W*\(|o(?:(
?:n(?:v(?:ert(?:_tz)?)?|cat(?:_ws)?|nection_id)|(?:mpres)?s|ercibility|alesce|t)\W*\(|llation\W*\(a))|d(?:(?:a(?:t(?:e(?:(?:_(add|forma
t|sub))?|diff)|abase)|y(name|ofmonth|ofweek|ofyear)?)|e(?:(?:s_(de|en)cryp|faul)t|grees|code)|ump)\W*\(|bms_pipe\.receive_message\b)|
(?:;\W*?\b(?:shutdown|drop)|\@\@version)\b|'(?:s(?:qloledb|a)|msdasql|dbo)'))" \
"phase:2,rev:'2.2.5',capture,t:none,t:urlDecodeUni,ctl:auditLogParts=+E,block,msg:'SQL Injection Attack',id:'950001',tag:'WEB
_ATTACK/SQL_INJECTION',tag:'WASCTC/WASC-19',tag:'OWASP_TOP_10/A1',tag:'OWASP_AppSensor/CIE1',tag:'PCI/6.5.2',logdata:'%{TX.0}',severi
ty:'2',setvar:'tx.msg=%{rule.msg}',setvar:tx.sql_injection_score=+%{tx.critical_anomaly_score},setvar:tx.anomaly_score=+%{tx.critical
_anomaly_score},setvar:tx.%{rule.id}-WEB_ATTACK/SQL_INJECTION-%{matched_var_name}=%{tx.0}"
```

What Didn't Work

- Manual IP Blocks
  - Hosts or networks
- Manual account closures/email address bans

- Attackers have many proxies or compromised hosts
- Email addresses are cheap and easy
- Manual methods are expensive
- Many attackers
- Attackers laugh at manual methods

IP Bans Almost Work

- IPv4 addresses are valuable
- Bans are easily automated
- Bans can time out automatically after a while
- Effective at discouraging attackers
  - Password brute-forcing essentially stopped
  - Spam sending attempts greatly diminished
    - Orders of magnitude
    - Less load on systems

SpamHaus

- Organization that fights Spam
  - Uses honeypots
- SBL
  - Verified Spam sources
- XBL
  - IPs of infected computers
  - You should be extremely worried if your systems show up in the XBL
- Read-only hubs if you're listed!

SpamHaus at Purdue

- Got a SpamHaus account for Purdue University (free, fairly easy)
- Found several infected IPs
  - Despite internal scanning and border IPS
- No false positives so far (1 month)
- New students coming in!
  - Waiting to see how many will bring infected computers?
- Highly recommended

IP Ban Problems

- IPv4 addresses are shared
  - HTTP Proxies
    - Asian universities
    - Hospitals
  - Wireless with address translation
    - Airports, shops
  - Satellite Internet
    - Africa
    - Countryside

- NAT and proxies make IP blocks problematic
  - Block legitimate, innocent users
- IP addresses get reassigned
- IPv6 addresses are too numerous to ban individually
  - Privacy extension
- Surprise:
  - Many hospital proxies listed in XBL as infected, despite HIPAA

Anti-Virus

- Scan all uploads
  - Reject infected files
- Rescan uploads periodically
  - In case newer definitions could catch something

- AV issues
  - Lag in signatures
    - Always catching yesterday's threat
  - False positives
  - Incomplete
    - False negatives
  - Bugs

Things that Could Help

- Accountability!
- More valuable accounts
  - Closing an account should hurt
- Automated ways to detect and block abusive behavior
  - Quickly
  - Precisely
- Behavior and reputation become important

- Banning solution for IPv6
  - Need to ban networks
  - Adaptive: match ban size to source size
  - Speed our adoption of IPv6
- Diligent proxy and NAT operators
  - Block attacks originating from your networks
    - Egress filtering
  - Provide recourse

“For me, trust is the availability of effective recourse” - Dan Geer

Possibilities

- Account Tiers
  - Newbie account with limited capabilities
    - Limited opportunity to cause harm
  - Grant capabilities based on
    - Accounts linked (Google, Facebook)
      - Shibboleth federated authentication
        - Transitive trust
    - Age of account
    - Participation
    - Vetting by someone else
  - Capability loss for bad behavior

Wishes

- Easy to use and customize web application firewall
- Banning software effective with IPv6
- No proxies/NAT/Sattelite networks protecting infected computers with innocent users
- Fewer infected computers
  - More people that monitored SpamHaus blacklists!

FYI

- Our lists of long duration bans are publicly available at
  - “Hub”/bans
    - Nanohub.org/bans
    - Hubzero.org/bans
    - Nees.org/bans