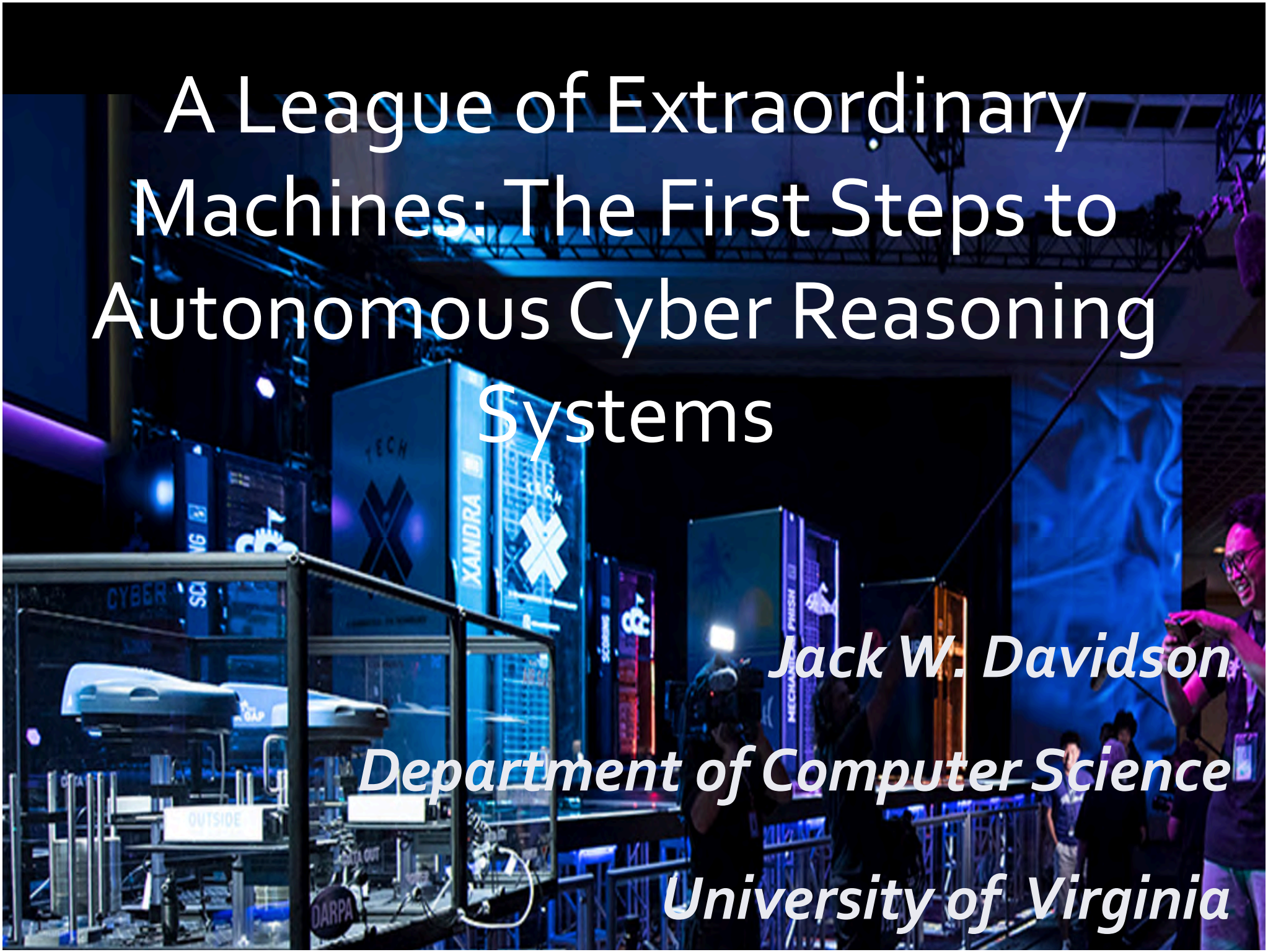
A League of Extraordinary Machines: The First Steps to Autonomous Cyber Reasoning Systems



Jack W. Davidson

Department of Computer Science

University of Virginia

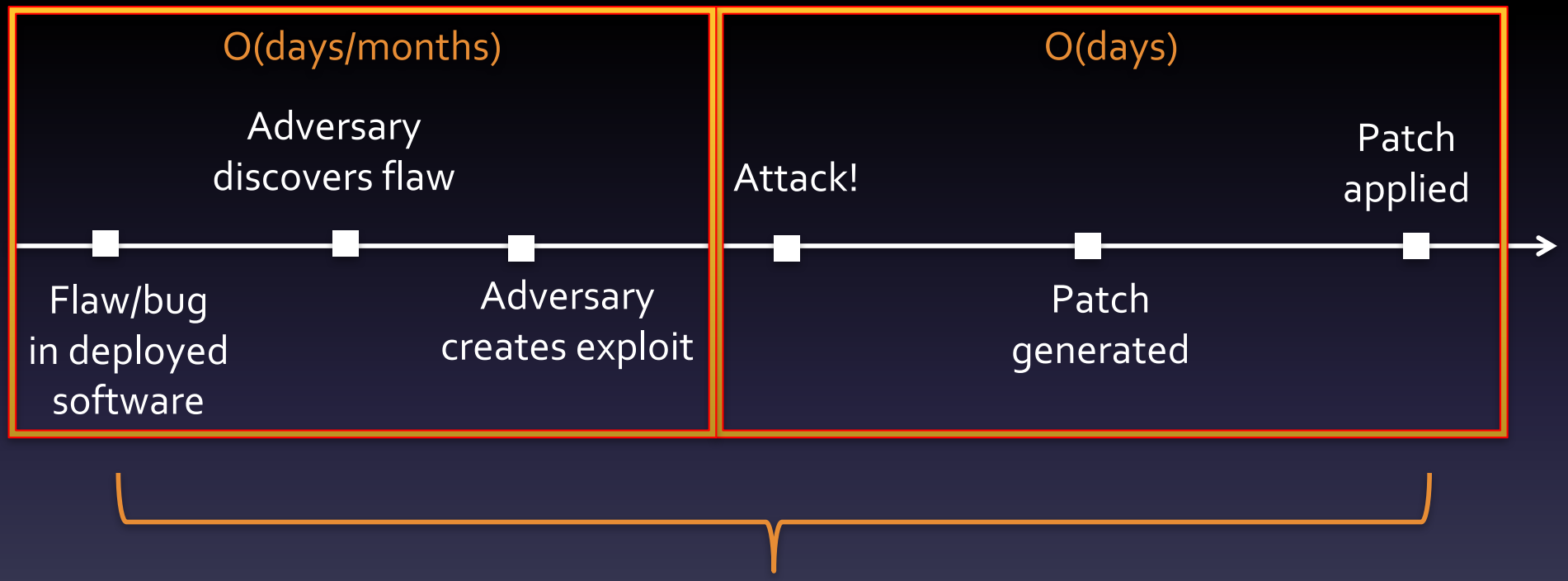
Cyber Grand Challenge

- International research competition to design and build a special-purpose "supercomputer" or cyber reasoning system that automatically discovers, confirms, and fixes software flaws in seconds, proactively preventing cyber intrusions
  - $2M first prize, $1M second prize, $750K third prize
- Challenge: Build an autonomous machine that can play capture the flag.

Why Autonomous Cyber Defense?

Internet of Things

Why?


O(days/months)
Adversary discovers flaw
Flaw/bug in deployed software
Adversary creates exploit
O(days)
Attack!
Patch generated
Patch applied
Window of vulnerability (days/months)

Cyber Capture the Flag

104 teams at start

28 teams in qualifying event

7 teams qualify for final event

Final event Las Vegas

June 2014

June 2015

trials

Aug 2016

Idaho (Jima)

Deep Red (Rubeus)

Disekt (Crspy)

ForAllSecure (Mayhem)

ShellPhish (Mechaphish)

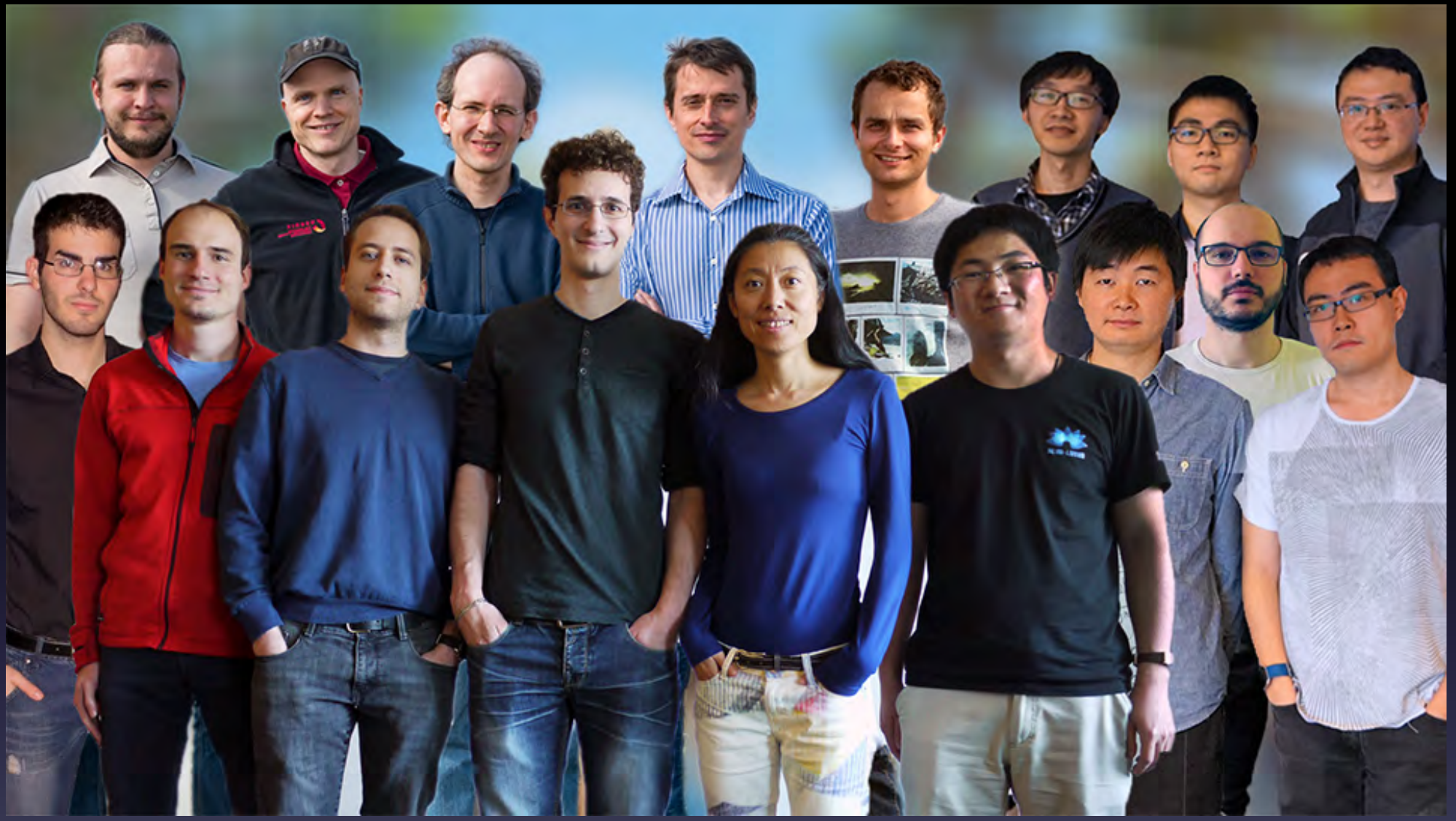
TechX (Xandra)

CodeJitsu (Galactica)

CSDS: University of Idaho
Machine: Jima

Deep Red: Raytheon Corporation
Machine Name: Rubius

Disekt: University of Georgia
Machine Name: CRSPY



Codejitsu: U. of California, Berkeley
Machine Name: Gallatica

For all Secure: Pittsburgh, PA
Machine Name: Mayhem

Shellphish: U. of California, Santa Barbara

Machine Name: Mechaphish

TechX: U. of Virginia & Grammatech, Inc.
Machine Name: Xandra

Cyber Grand Challenge
Research Challenges

- High-precision static and dynamic analysis of previously unseen binary code
- Automatic identification of vulnerabilities in binaries
- Create proofs of vulnerabilities
- Automatic creation and application of patches to mitigate vulnerabilities without damaging software
- Operate at cyber speed: Identify vulnerabilities and patch within seconds or minutes
- No human in the loop: fully autonomous

Evaluation (Proof of Vulnerability)

Type 1 (subvert control flow)

- Control 20+ bits of instruction pointer on crash
- Control 20+ bits of general purpose register

Type 2 (information leakage)

- Leak 4 bytes from flag page, a memory-mapped page at known location filled with random data

Evaluation = 1 + n/6 (n <= 6)

Security (defense)

Replace binary or install firewall rule:
1 round penalty

Security = {
- 1 if **any** competitor throws successful POV
- 2 if **no** competitor throws successful POV

Scoring

100 x

availability x **security** x **evaluation**

0..1 1 or 2 1+n/6 (n<=6)

Range = [0..400]

XANDRA ARCHITECTURE

Xandra

Hardware/Software

64 nodes

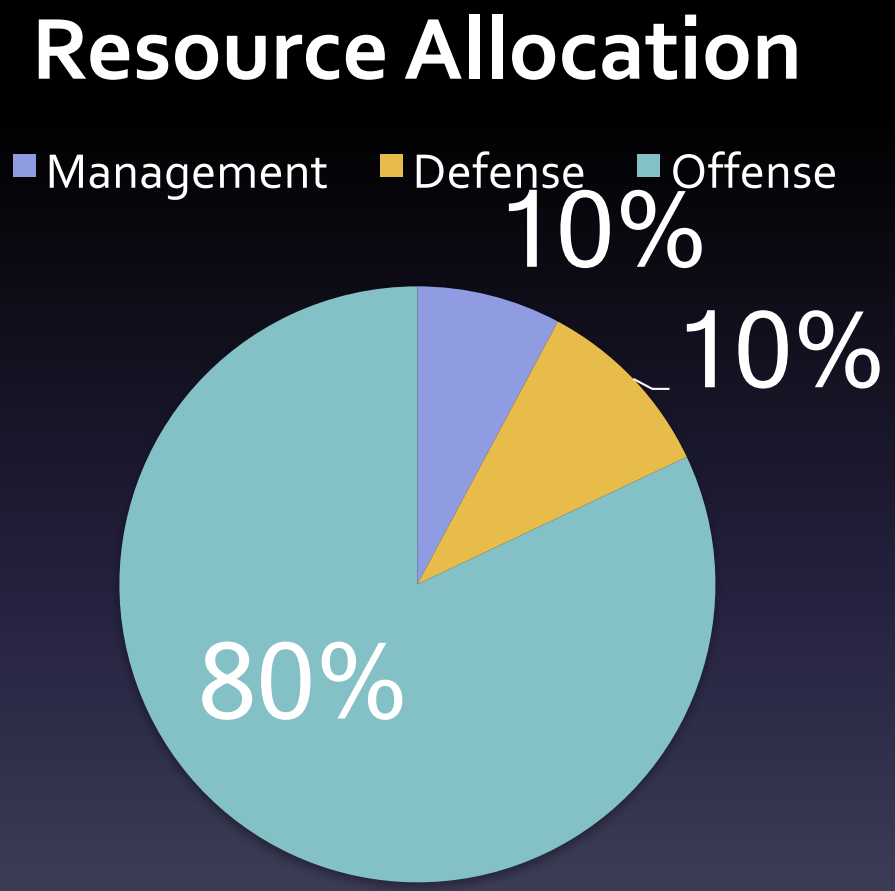
1280 cores, 2560 vCPUs

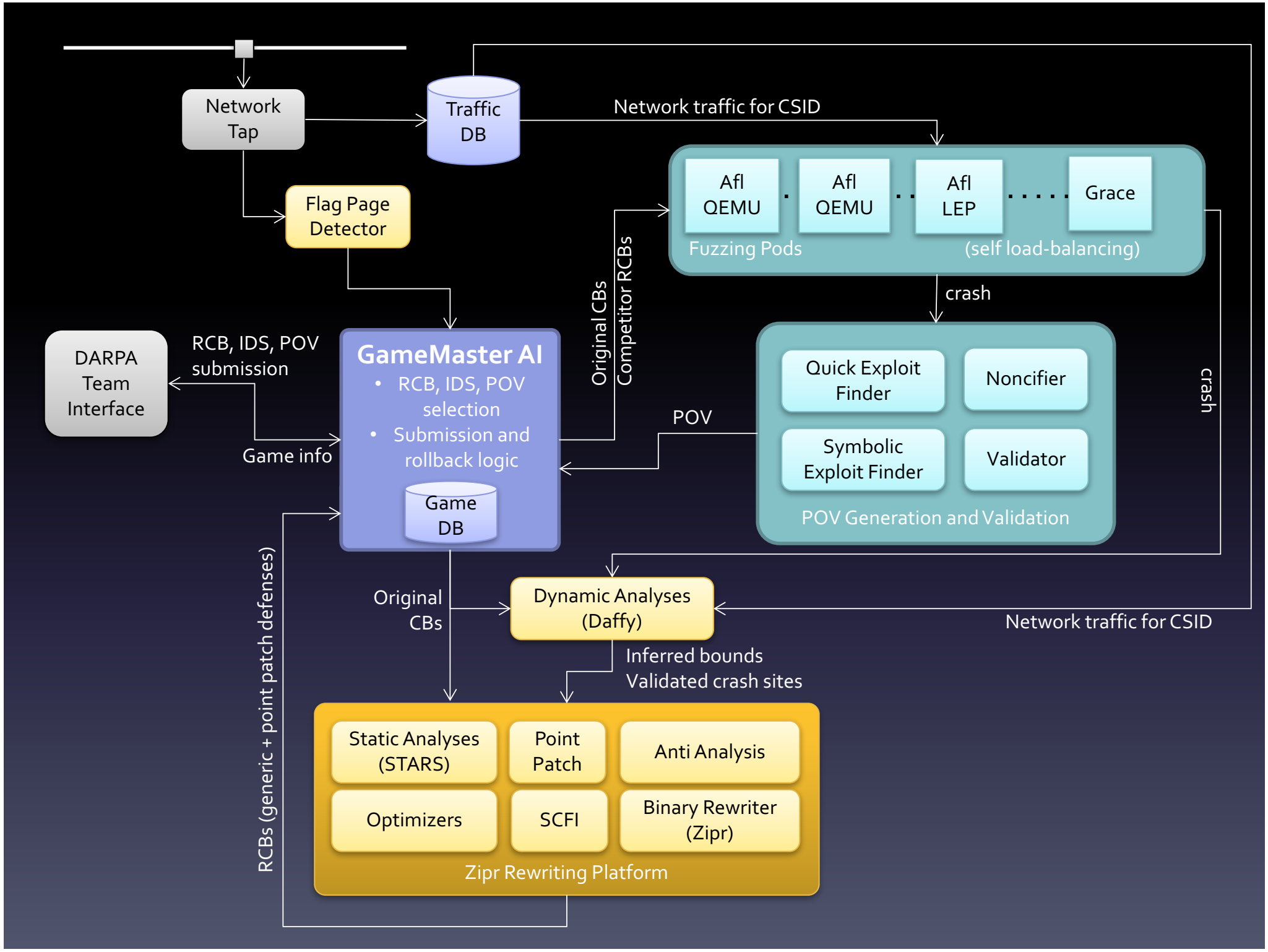
16 TB RAM

128 TB Storage

Openstack, Ubuntu 14.04

- OpenStack cloud infrastructure
- Bag-of-tasks architecture
  - Naturally self-load balancing
  - Naturally fault-tolerant
- Segregation trusted/untrusted workers




Network Tap
Traffic DB
Network traffic for CSID
Flag Page Detector
Afl QEMU
Afl QEMU
Afl LEP
Grace
Fuzzing Pods
(self load-balancing)
crash
DARPA Team Interface
RCB, IDS, POV submission
GameMaster AI
• RCB, IDS, POV selection
• Submission and rollback logic
Game DB
Original CBs
Competitor RCBs
Game info
POV
Quick Exploit Finder
Noncifier
Symbolic Exploit Finder
Validator
POV Generation and Validation
crash
Original CBs
Dynamic Analyses (Daffy)
Network traffic for CSID
Inferred bounds
Validated crash sites
RCBs (generic + point patch defenses)
Static Analyses (STARS)
Point Patch
Anti Analysis
Optimizers
SCFI
Binary Rewriter (Zipr)
Zipr Rewriting Platform

Xandra Defenses

- Block-level Instruction location (BILR)
- Selective Control-flow Integrity (SCFI)
- Daffy and Point-patching
- Binary optimization
- Anti-analysis techniques
- Network defenses

Xandra SCFI

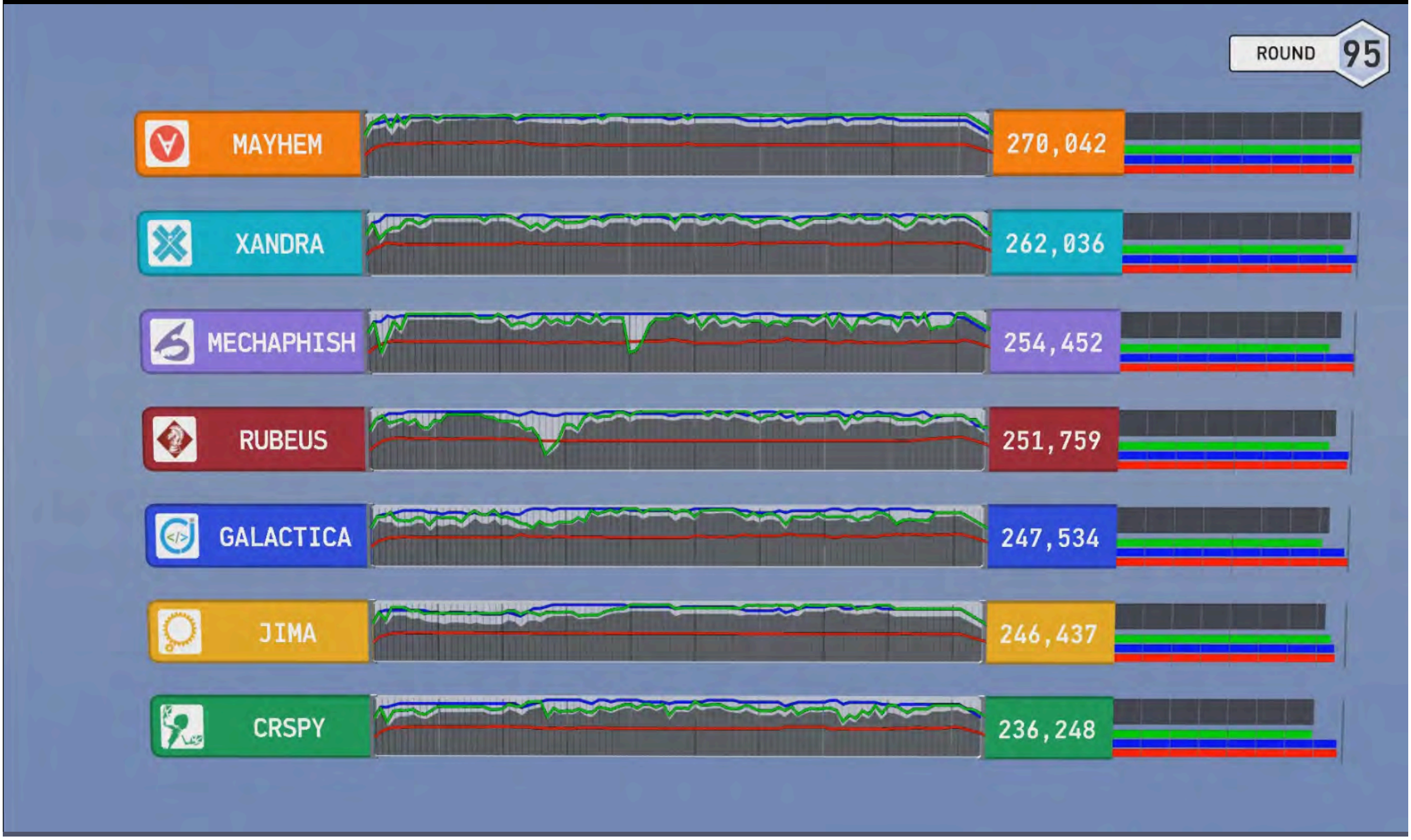
- Coarse-grained: All indirect control-flow transfers—targets of indirect jumps, calls and returns—belong to the same target class
- Use formal methods to prove certain indirect branches safe and do not protect

```
(1) ... ; at call to foo():
(2) call foo
(3) nop ; 1-byte executable nonce 0x90

(4) ... ; at return from foo():
(5) and [esp], 0x7FFFFFFF ; clamp
(6) mov ecx, DWORD [esp] ;
(7) cmp BYTE [ecx], 0x90 ; verify nonce
(8) jne _terminate
(9) ret
```


CGC

Final Scoreboard



Scoring Breakdown

CRS	Security (defense)	Evaluation (offense)	Availability (func, overhead)
1. Mayhem	#6	#6	#1
2. Xandra	**#1**	**#4**	**#2**
3. Mechaphish	#2	#1	#5
4. Rubeus	#3	#3	#4
5. Galactica	#4	#2	#6
6. Jima	#7	#7	#3
7. Crspy	#5	#5	#7

Only 1 instance (1 challenge set for 1 round) where a competitor was able to bypass Xandra's defenses

Defensive Gains

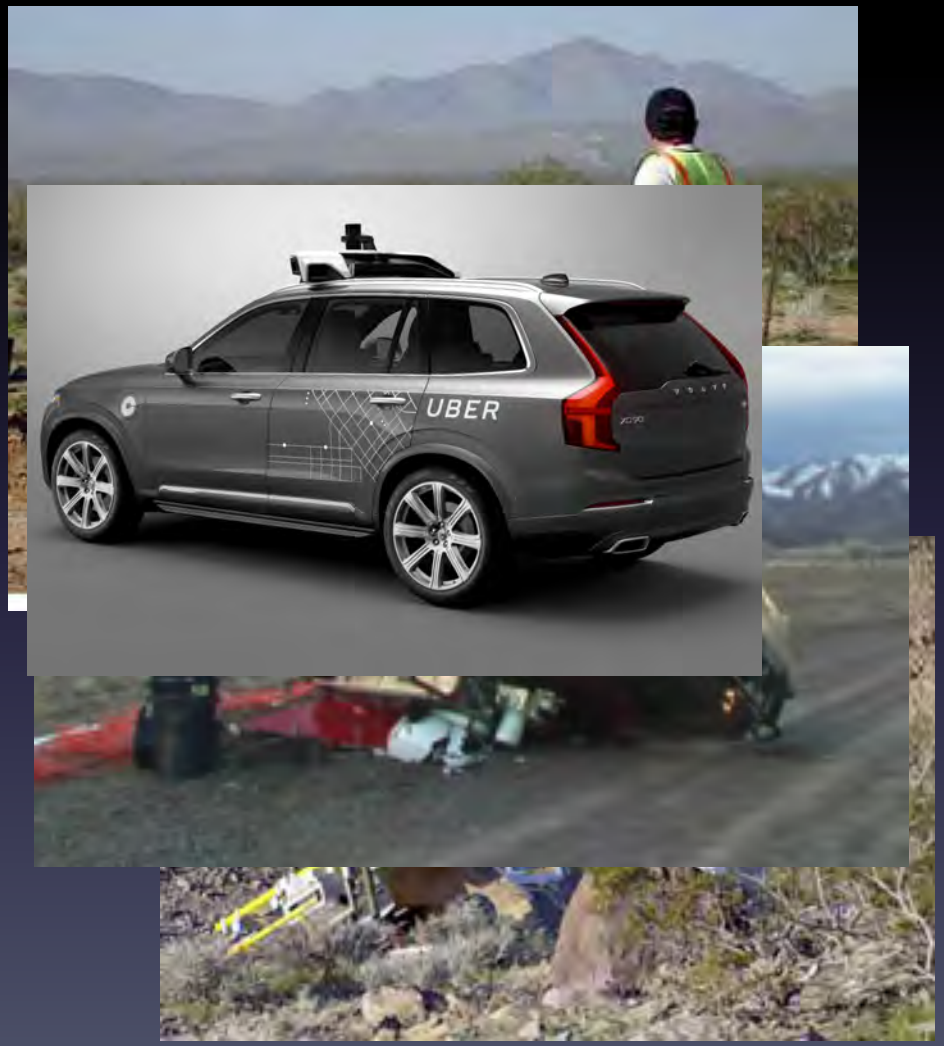
CRS	Never POVed	POVed	Defensive Gains
1. Mayhem	(477)	8,849	8,372
2. Xandra	**(13,441)**	**15,071**	**1,630**
3. Mechaphish	(25,308)	13,162	(12,146)
4. Rubeus	(10,901)	473	(10,429)
5. Galactica	(25,385)	8,188	(17,197)
6. Jima	(10,903)	244	(10,659)
7. Crspy	(27,971)	3,280	(24,690)

DEFCON 24 CTF

Team	Final Score
PPP	113555
b1oop	98891
DEFKOR	97468
HITCON	93539
KaisHack GoN	91331
LC↯BC	84412
Eat Sleep Pwn Repeat	80859
binja	80812
pasten	78518
Shellphish	78044
9447	77722
Dragon Sector	75320
!SpamAndHex	73993
侍	73368
Mayhem	72047

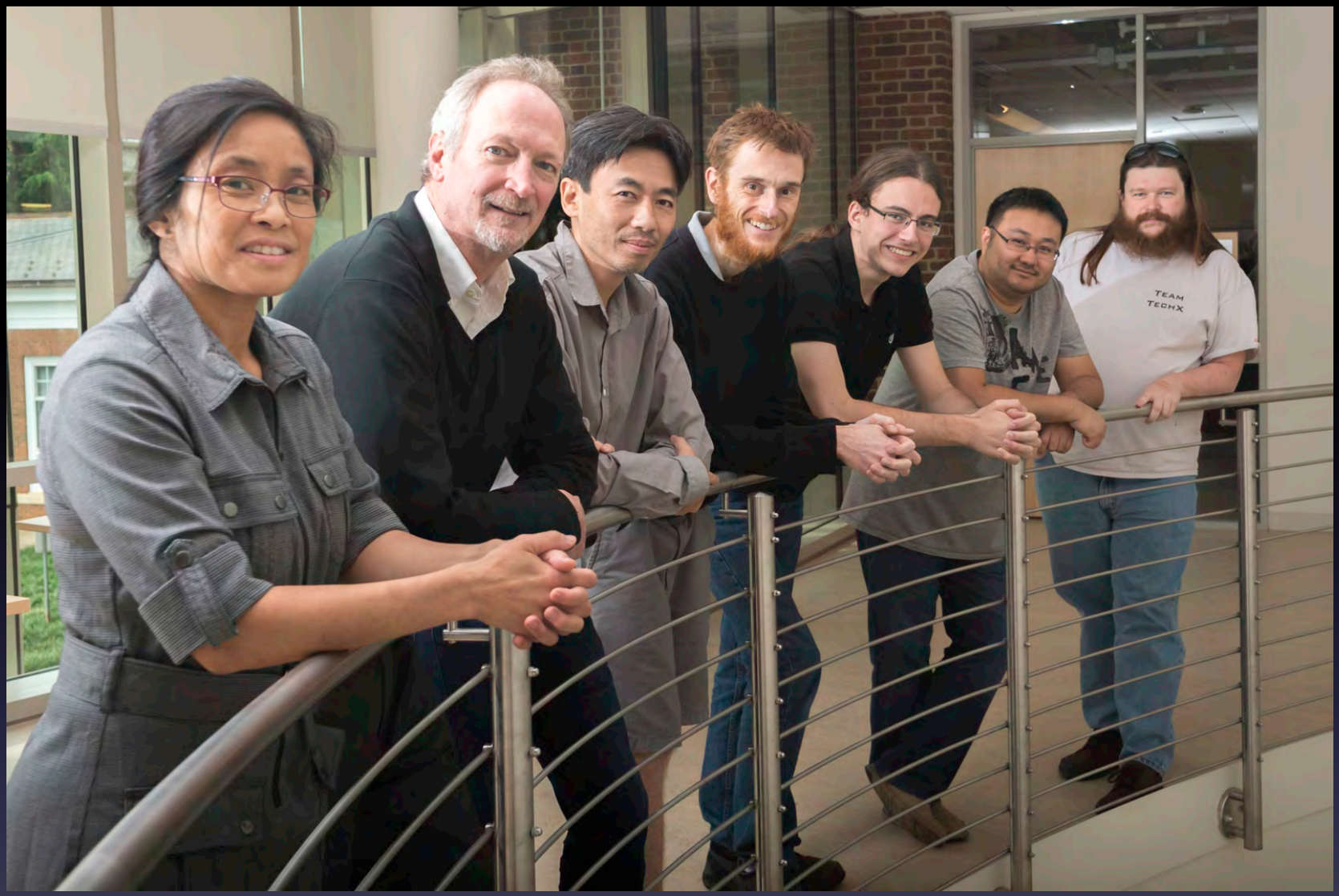
Was Cyber Grand Challenge a Success?

- Demonstrated that fully automated cyber defense is achievable
- Systems were able to identify and patch critical vulnerabilities in under five minutes: Heartbleed, Slammer, sendmail
- Missed many vulnerabilities
- Systems easily beaten by human teams
- Many research challenges ahead!

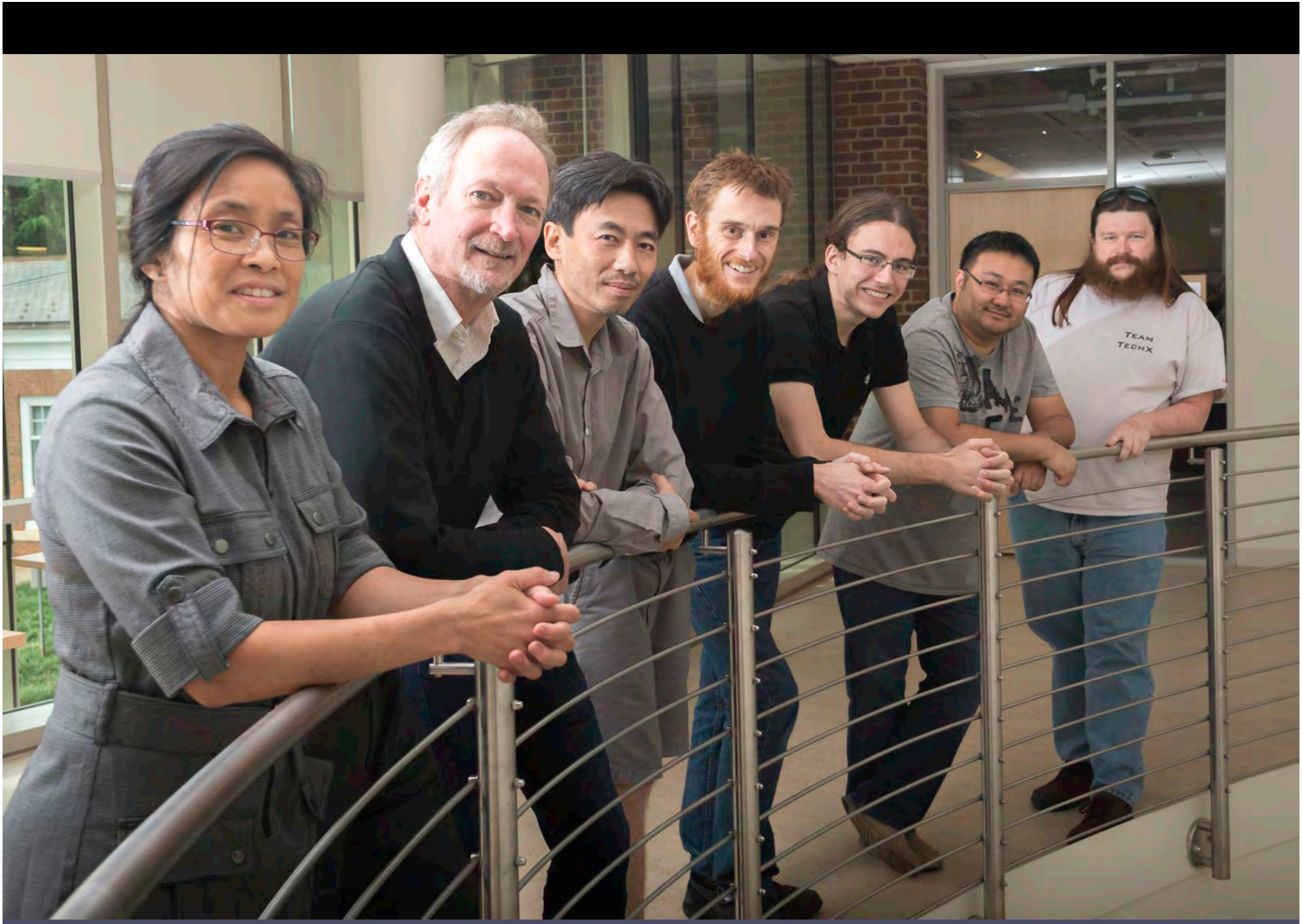


Concluding Thoughts

- Fully autonomous systems will soon be commonplace: smart cities, smart homes, autonomous vehicles, assistive robots, etc.
- The impact of these systems on society will be profound
- We, as computer scientists and engineers, must:
  - Understand their impact on society
  - Understand the risks and consequences of attacks on these systems
  - Ensure these systems operate as intended and the data they collect and process is secure from improper use
- Overall, I see a bright future ahead!



University of Virginia TechX Team


TEAM
TECHX