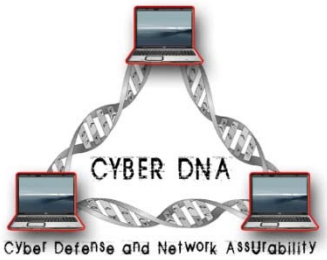


Formal Configuration Analytics for Provable & Measurable Security

Ehab Al-Shaer

Cyber Defense & Network Assurability (CyberDNA) Center
Department of SIS, College of Computing and Informatics
University of North Carolina Charlotte
www.cyberDNA.uncc.edu/~ehab

NSF US/Mideast Workshop on Trustworthiness in Emerging Distributed Systems and Networks
June 4-6, 2012

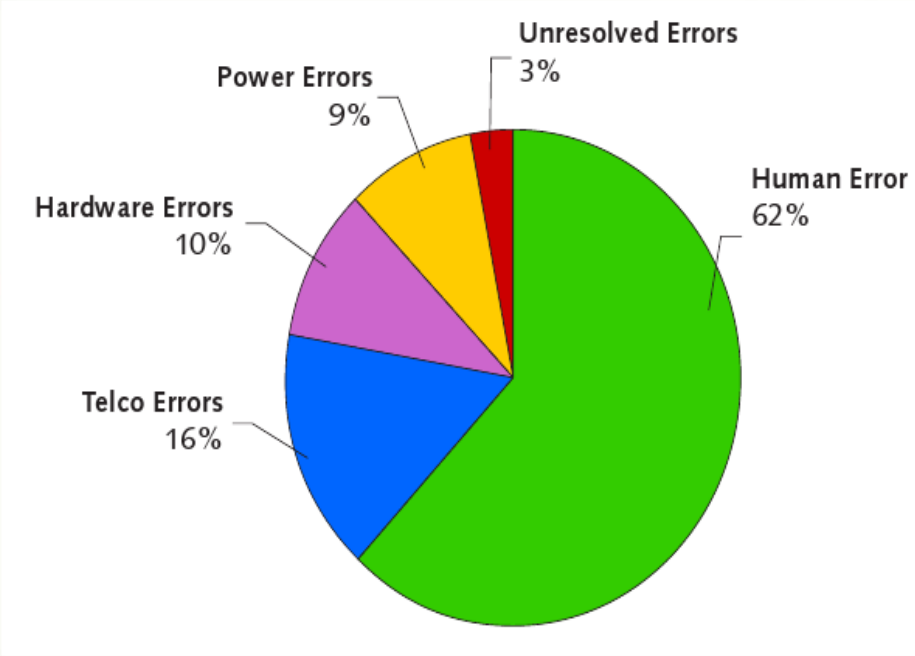
UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

- Enforcing security and privacy objectives (identity management, intrusion detection cyber defense) is dependent on correct (and optimizing) security configurations.
- Cyber is a complex system of heterogonous interconnect systems:
  - End hosts configuration (services, access controls, VM, registries etc)
  - Firewalls
  - IPSec
  - IDS
  - Proxies
  - AP
  - Mobile and pervasive devices
  - Internet of things: sensor/actuators
  - Smart devices for smart grid and others
  - etc
- Assessing security requires holistic characterization of system (end-to-end) behavior considering the interdependences/interrelation between different components and configuration
- Formal analytics techniques are needed for provable configuration security

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics



"Eighty percent of IT budgets is used to maintain the status quo.", Kerravala, Zeus. "As the Value of Enterprise Networks Escalates, So Does the Need for Configuration Management." *The Yankee Group* January 2004 [2].

"Most of network outages are caused by operators errors rather than equipment failure.", Z. Kerravala. Configuration Management Delivers Business Resiliency. The Yankee Group, November 2002.

- "It is estimated that configuration errors enable **65%** of cyber attacks and cause **62%** of infrastructure downtime", Network World, July 2006.
- *Recent surveys show Configuration errors are a large portion of operator errors which are in turn the largest contributor to failures and repair time [1].*
- *"Management of ACLs was the most critical missing or limited feature, Arbor Networks' Worldwide Infrastructure Security Report, Sept 2007.*

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

- December 2008 report from Center for Strategic and International Studies "Securing Cyberspace for the 44th Presidency" states that "**inappropriate or incorrect security configurations were responsible for 80% of Air Force vulnerabilities**"
- May 2008 report from Juniper Networks "What is Behind Network Downtime?" states that "**human factors [are] responsible for 50 to 80 percent of network device outages**".
- BT/Gartner[3] has estimated that **65% of cyber-attacks exploit systems with vulnerabilities introduced by configuration errors. The Yankee Group[4] has noted that configuration errors cause 62% of network downtime**.
- A 2009 report[5] by BT and Huawei discusses how service outages caused by "**the human factor" themselves cause more than 30% of network outages, "a major concern for carriers and causes big revenue-loss.**

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

Misconfigured networks create huge security risks

Bill Brenner, Senior News Writer

Published: 05 Mar 2008

There's a perpetual buzz around software flaws and exploits researchers disclose daily, but security experts say it often distracts IT pros from a growing and more serious problem --

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

Making the case for network security configuration management

Tom Bowers, Contributor

Let's be realistic: The discussion of network security configuration management doesn't make security pros excited to jump out of bed in the morning. It's simply one of those tasks that must be done.

The problem is it should be top-of-mind, as a failure to properly manage network security configurations can be a career-ender, or as the security community calls it, a "Resume Producing Event (RPE)."

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

IIS configuration error leads to increased threat, Microsoft says

SearchSecurity.com Staff

Published: 04 Jan 2010

Microsoft said an Internet Information Services (IIS) parsing extension issue,which could lead to a vulnerable system, is not a flaw that can be patched, but an IIS configuration error that can be avoided by following best practices.

Microsoft IIS best practices:

IIS 6.0 security best practices: Microsoft TechNet document outlines best practices for configuring the Web server.

Microsoft updates:

Dec. - Microsoft gives Internet Explorer a major security overhaul: The final regular Microsoft update of 2009 repairs five critical vulnerabilities in IE and blocks

The software giant issued an update on its blog last week, giving links outlining best practices for configuring the IIS Web server. A security expert warned last week about the discovery of a parsing extension vulnerability that could be exploited to pass malicious code and ultimately gain control of the Web server. The issue was described as an error in the way IIS 6 handles semicolons in URLs.

But Microsoft's Christopher Budd explained on the company's Security Response Center blog that the issue is a IIS configuration error that could lead to a vulnerable system. The out-of-the-box, default configuration will not

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

Misconfiguration issues could have contributed to Hannaford breach

Robert Westervelt, News Editor

Published: 19 Mar 2008

The fallout over the data breach at Hannaford Bros. continued Wednesday, as Massachusetts officials suggested the supermarket chain was too slow in disclosing the incident and one of the retailer's security vendors went on the defensive.

This demonstrates that there are a lot more targeted attacks out there and the targeted attacks have a high monetary risk. David Precopio, vice president of marketing and business

Officials suggested in published reports that under state law, Hannaford should have notified the Massachusetts Office of Consumer Affairs and Business Regulation as soon as the company became aware of it. As of Wednesday afternoon, the consumer affairs office had yet to receive the official notification. The law took effect last year in the wake of the massive data breach at Framingham, Mass.-based TJX Companies Inc.

The Maine-based supermarket chain revealed Tuesday that it first detected something amiss three weeks ago but that it stalled its

UNC CHARLOTTE
College of Computing and Informatics

Why Configuration Analytics

- **Tufin Technologies study (2011)** : "Nearly **85 percent** of network administrators in the 2011 Firewall Management report said half of their firewall rule changes need to be fixed because they were configured incorrectly"
  http://www.eweek.com/c/a/Security/Majority-of-Firewall-Rules-are-Improperly-Configured-Managed-Survey-Finds-388413/
- **Gartner Research study (2008)**: "More than **99% of firewall security breaches** are caused by configuration mistakes"
  http://www.techdata.com/techsolutions/networking/files/june2010/gartner%20firewall%20page%207%20qa_is_it_more_secure_to_use__160362.pdf
- **Tufin Technologies study (2011)**: "a November 2011 survey from Tufin of 100 firewall managers revealed that **only 1.3%** of configuration changes that cause network downtime or pose a security breach **are identified during the quarterly audit**"
  http://www.tufin.com/blog/2012/01/24/network-security-101-automating-for-continuous-compliance/

Science of Configuration Analytics & Automation

- ***Science of Configuration (SoC)*** *is to constitute a scientific methodology for (1) creating and validating hypotheses* (properties) about the **global system behavior** based on its components' configurations or logs, and (2) synthesizing **configurations of composite components** that can provably satisfy global system properties deterministically or probabilistically.
- **Configuration vs. s/w or h/w**
- **Configuration Analytic** (Bottom-up): modeling, verification, repair, optimization, measuring/metrics
- **Configuration Automation** (top-donw): integration/unification, synthesis, planning, tuning /autonomics

Modeling ACL Configuration Using BDDs

- An ACL policy is a sequence of filtering rules that determine the appropriate action to take for any incoming packets: $P = R1, R2, R3, .., Rn$
- Each rule can be written in the form:

$$R_i := C_i \rightsquigarrow a_i$$

where C_i is the constraint on the filtering fields that must be satisfied in order to trigger the action a_i

- The condition C_i can be represented as a Boolean expression of the filtering fields $f_1, f_2, \ldots, f_k$ as follows:

$$C_i = fv_1 \wedge fv_2 \wedge \cdots \wedge fv_k$$

where each fv_j expresses a set of matching field values for field f_j in rule R_i. Thus, we can formally describe a ACL policy as:

EA1

$$P_a = (C_1 \wedge b_1) \vee (\neg C_1 \wedge C_2 \wedge b_2) \ldots \vee (\neg C_1 \wedge \neg C_2 \ldots \neg C_{i-1} \wedge C_i \wedge b_i)$$

rule1 rule2 rule_n

$$\text{where } b_i = \begin{cases} 1 \text{ if } action_i = a \\ 0 \text{ if } action_i \neq a \end{cases}$$

Slide 11

EA1 policy is a disjunction expression of set of conjuctive terms where each term i represents the firing condition for rule i

since action must be binary here ==> I need an expression for each action exists in the policy

Ehab Al-Shaer, 9/21/2005

Concise Formalization

- Single-trigger policy **is an access policy where only one action is triggered for a given packet.** C_i is the 1st match leads to action *a*

$$P_a = \bigvee_{i \in index(a)} (\neg C_1 \wedge \neg C_2 \ldots \neg C_{i-1} \wedge C_i)$$

$$P_a = \bigvee_{i \in index(a)} \bigwedge_{j=1}^{i-1} \neg C_j \wedge C_i$$

- Multiple-trigger policy **is an access policy where multiple different actions may be triggered for the same packet.** C_i is any match leads to action *a*

$$P_a = \bigvee_{i \in index(a)} C_i$$

where $index(a) = \{i \mid R_i = C_i \rightsquigarrow a\}$

ESS2

ESS1

Formalization – The Basic Model

- The network is modeled as a state machine
  - each state determined by the packet header information and packet location on the network
    - States = Locations X Packets
  - The ***characterization function*** to encode the state of the network in the basic model (abstracting payload)

$$\sigma : \mathrm{IP_s} \times \mathrm{port_s} \times \mathrm{IP_d} \times \mathrm{port_d} \times \mathrm{loc} \rightarrow \{\mathbf{true}, \mathbf{false}\}$$

$\mathrm{IP_s}$	the 32-bit source IP address
$\mathrm{port_s}$	the 16-bit source port number
$\mathrm{IP_d}$	the 32-bit destination IP address
$\mathrm{port_d}$	the 16-bit destination port number
loc	the 32-bit IP address of the device currently processing the packet

Slide 13

ESS1 - if the function is true means this pkt with this specific header exist in this specif locaton

- if there are 5 pkts in the network then the fucntion will have 5 statifying assignments

- I.e., every statisfying assignmetn for this fucntion represnerts a at least one packet with specific header exists on a specific location

Ehab Al-Shaer, 4/9/2010

ESS2 why location?

- allows for global heterogneious devices analysis
- more scalable than rul abstraction
- allows for investigating devices-specific problmes
- faster for rebuilding in dynamci update

Ehab Al-Shaer, 4/9/2010

Formalization – The Basic Model

- Network devices are modeled based on the **packet *matching semantic*** and ***packet transformation***
  - Each rule consists of a condition (Ci) and an action (a): Ci➔a
  - Policy are set of rules matched sequentially with single- or multi-trigger actions
  - Firewall (single trigger) policy encoding using BDD

$$P_a = \bigvee_{i \in index(a)} (\neg C_1 \wedge \neg C_2 \ldots \neg C_{i-1} \wedge C_i)$$

$$= \bigvee_{i \in index(a)} \bigwedge_{j=1}^{i-1} \neg C_j \wedge C_i$$

- **Transformation:**
  - if a pkt *state* matches the rule *condition*, the Action can change the packet location and possibly the headers ➔ means change over the bits of the *state*
- Transition relation is ***characterization function*** as follows:
  - **t**: (Curr_pkt x Curr_loc)x (New_pkt x New_loc) → {true, false}
  - Device Model ϕ = ***loc*** $\wedge$ ***Match_Condition*** $\wedge$ **t → {true, false}**

Formalization – The Basic Model

- Global Transitions relation of the entire network:

$$T = \bigvee_{i \in devices} \Phi_{device_i}$$

- Variables
  - Locations is every place that can describe packet position: firewall, router, IPSec device, or application layer service, etc.
  - We allow Location to be different than IPsrc for spoofing
  - There are two versions of each variable: current and new state.
- Each property and field describing the state (i.e., location IP; packet properties: src/dst IP; port, proto, transformation, etc) is represented by bits, according to its size.
- These variables are used via a symbolic representation using Ordered Binary Decision Diagrams.
- Model Checking and CTL are used to answer the queries posed by the administrator.

ConfigChecker: Global Configuration Analytics in a Box

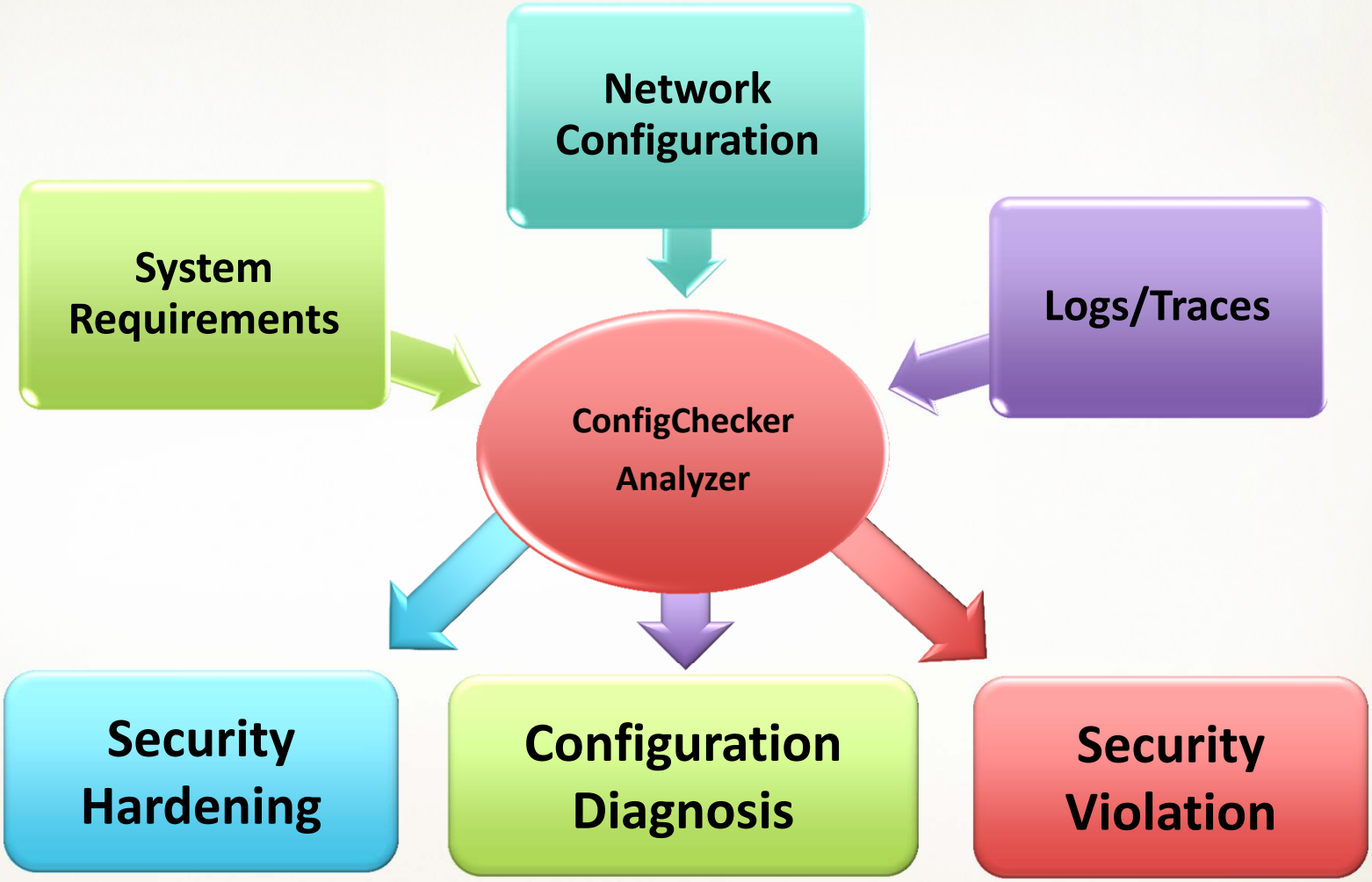
UNC CHARLOTTE
College of Computing and Informatics

CONFIGCHECKER: What and How?

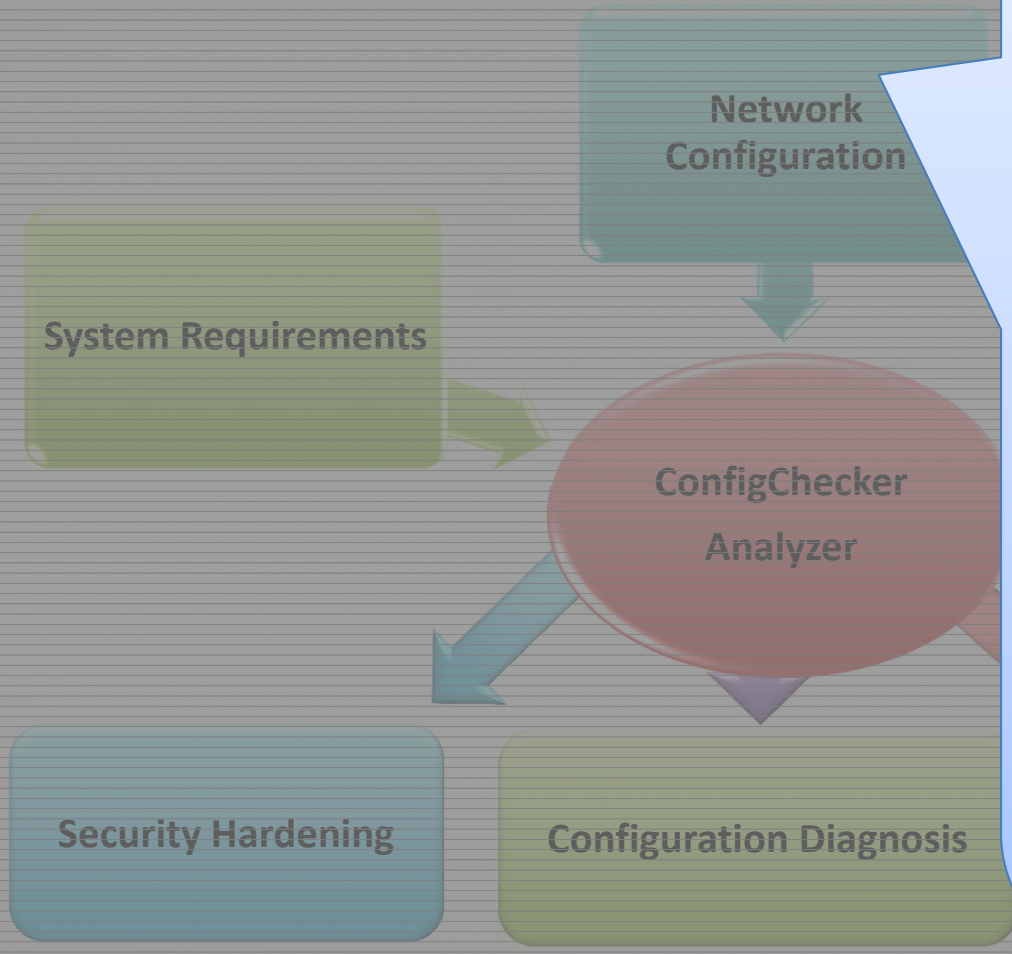
- **What is ConfigChecker?**
  - Automated security analytic tool for global network configuration verification, diagnosis, repair and hardening
  - ConfigChecker Engine is based on symbolic model checker using BDDs and SAT tools.
  - ConfigChecker allows for abstraction and composition; the entire network configuration represented in a Boolean formula
  - Bottom-up analysis: given security and risk requirements, find what is non-compliant in the existing configuration.
- **Why ConfigChecker is Unique?**
  - Global end-to-end analytics that includes routing, firewalls, NAT, IPSec/VPN, IDS, multicast, proxies, wireless AP, VMs, smart phone.
  - Provable analytics (not simulation)
  - Scalable (more than **4000s** of devices, and 6 **millions** rules)
  - Provides an expressive Temporal Logic-based interface languages
  - Integrate configuration and (host/network) log analysis

UNC CHARLOTTE
College of Computing and Informatics

WHAT IS CONFIGCHECKER?

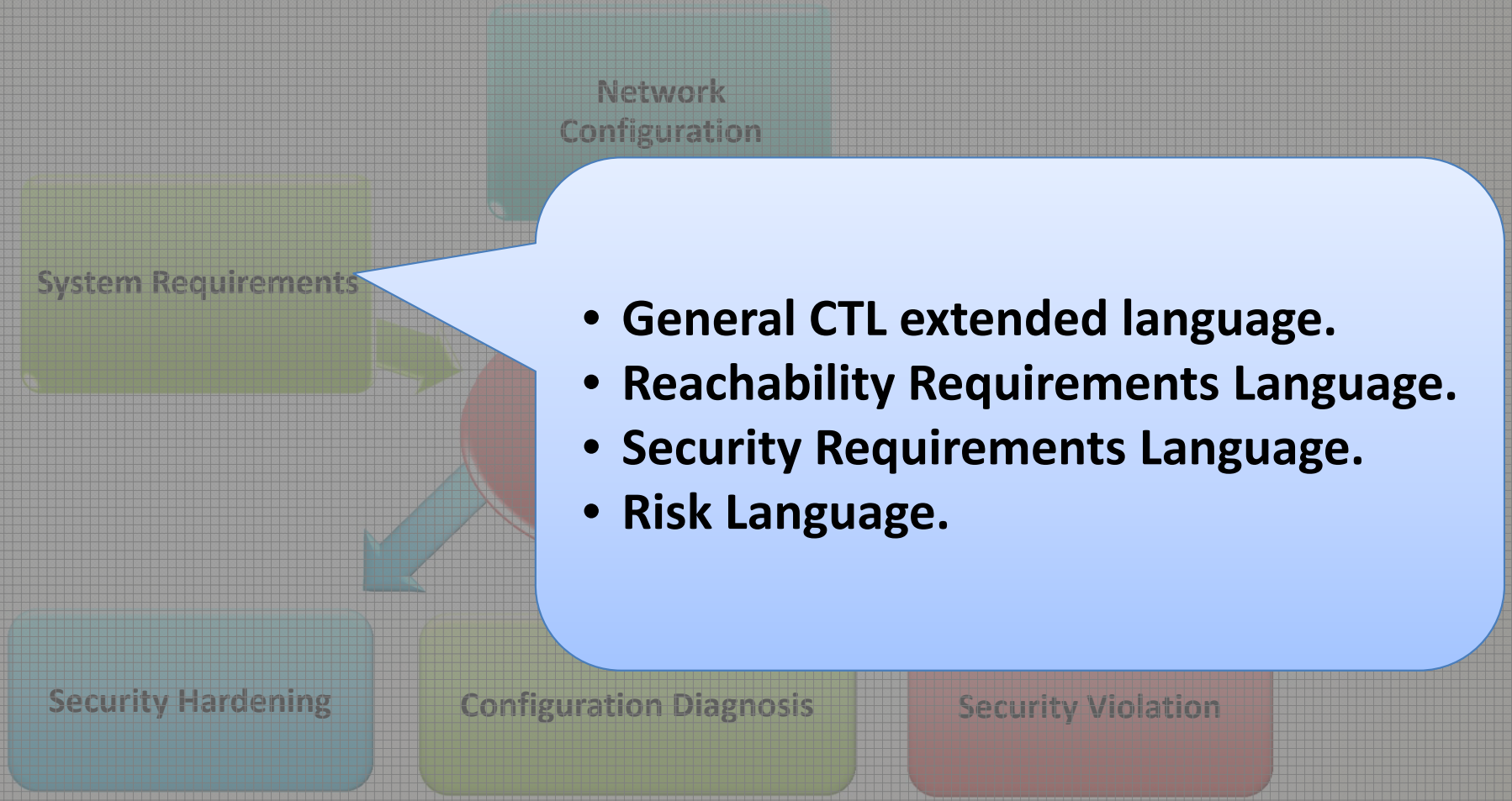


Architecture



- **Network Configuration**
  - Routing Tables
  - Firewalls ACL
  - NAT Tables.
  - IPSec Transformation.
  - IDS
  - Wireless AccessPoints
- **Host Configuration**
  - Gateway and subnet mask.
  - Services.
  - Applications.
  - Mobility

Architecture



- **General CTL extended language.**
- **Reachability Requirements Language.**
- **Security Requirements Language.**
- **Risk Language.**

UNC CHARLOTTE
College of Computing and Informatics

EXAMPLES OF CONFIGCHECKER ANALYTICS CAPABILITIES

- **Reachability Analysis:** Is there any service unreachable due to routing or firewall misconfigurations or potential route failures?
- **Security Analysis:** Is there any violation to a the end-to-end access control with required credentials?
- **Consistency Analysis:** Is there any *conflict* in policy actions or packet transformations in the network?
- **Risk Analysis:** Is there any violation to Risk-based access control policies defined based vulnerabilities, exposure and impact

UNC CHARLOTTE
College of Computing and Informatics

CONFIGCHECKER REACHABILITY ANALYTICS

- **Reachability Analysis:** Is there any service unreachable due to routing or firewall misconfigurations or potential route failures?
  - Proof of Routing Completeness: Considering services interdependency and access control rules, all network services are reachable by authorized users/customers.
    - E.g., Who ever access SQL server should also access authentication server before.
  - Proof of Routing Soundness: every entry in FIB is a valid forward entry.
    - No bogus routing bogus entries, no black holes, no routing loops.
    - End-to-end reachability does not violate access control rules (e.g., No file upload To Internet from local machines even using WiFi or Bluetooth.)
  - Routing Resiliency
    - Given link or router failure scenario, is there route redundancy? How many? What is the quality (number of hops) of each route? Is there disjoint of degree X redundant paths?
    - What firewalls changes are needed in order to allow for (1) at least 2 redundant routes, (2) disjoint for at least 50%, and (3) route length is less than x number of hops.
  - Consistency checking between firewalls and routers
  - Network and Application-level access control compliance: Every DB user-level access controls is implemented in network access control (routing and firewalls).
    - If the user x has access to a DB table in server y then routing and firewall must allow machine of x to access machine y.

UNC CHARLOTTE
College of Computing and Informatics

CONFIGCHECKER CONSISTENCY ANALYTICS

- **Consistency Analysis:** Is there any *conflict* in policy actions or packet transformations in the network?
  - Intra-policy Rule Anomaly Detection:
    - Is there any conflict between ANY two rules (or more) in the same device, e.g., firewalls?
    - Flow Shadowing and spuriousness
  - Inter-policy Rule Anomaly Detection:
    - Is there any conflict between ANY two similar devices or packet transformation in the network?
    - Upstream firewall accepts but downstream denies (spuriousness) or vice versa (shadowing)
  - Inter-device Anomaly Detection
    - Is there any conflict between ANY two similar devices or packet transformation in the network?
    - Packets should be inspected before encrypted
  - How to identify complete, partial and conditional action conflicts?
  - What are the most effective (minimum) fixes to resolve conflict?

UNC CHARLOTTE
College of Computing and Informatics

CONFIGCHECKER SECURITY ANALYTICS

- **Security Analysis:** Is there any violation to a the end-to-end access control with required credentials?
  - Are access control enforced across all paths? Is there any back door or hidden tunnels due to misconfiguration or failures?
    - Considering various communication means such as cyber, WiFi, cell, Bluetooth using many wired and wireless devices, the potential of access control violation due to misconfiguring is significant.
  - Trusted Path: are the required end-to-end credentials enforced?
    - Broken Tunnels (Encrypted data appears as plain text).
  - Least privilege enforcement
    - Can someone upload to the internet if he has access to sensitive SQL server?
  - Back doors through handheld devices (Bluetooth attacks).

UNC CHARLOTTE
College of Computing and Informatics

ConfigChecker Risk analytics

- Access control are defined based on *Risk* and *Impact* values
- Networks servers are assigned subjectively *Risk* and *Impact* , *where*
- *Risk = f (potential vulnerability, Exposure)*
  - Potential vulnerability: CVSS score form NVD
  - Exposure: how much this host is reachable directly and indirectly
  - For example, high Risk = CVSS > 50 and high-exposure
- **Examples:**
  - High risk host should not contact High Impact
  - Medium risk can only contact High Impact iff the traffic is encrypted
  - Low risk hosts can access medium impact hosts iff traffic is deeply inspected
- **Risk Analysis:** Is there any violation to Risk-based access control policies defined based vulnerabilities, exposure and impact
  - Is any user in high risk environment can not access classified data?
  - Is each traffic inspected before arriving servers of high impact?
  - Is every data passing through a high risk domain encrypted?

UNC CHARLOTTE
College of Computing and Informatics

Reachability Requirements Language

CanReach(Src, SrcProfile, Dest, DestProfile)

- IP, M
- Any

- Unique ID.
- Running services (server)
- Running Applications (client).
- Connectivity Capabilities (Wire, Wireless, Bluetooth, GPS, Cellular).
- Assets, Vulnerability, OS, etc.

UNC CHARLOTTE
College of Computing and Informatics

Reachability Requirements Language

- Linking different rules with logical operators
  - AND
  - OR
  - NOT
  - IF-THEN (Implies).

R1: CanReach($H_{1,}$ _, SQLServer, _)
R2: CanReach($H_{1,}$ _, _, Internal_DNS_Profile)
R3: CanReach($H_{1,}$ _, _, Internal_Kerberos)

R4 : IF R1 THEN (R2 AND R3)

UNC CHARLOTTE
College of Computing and Informatics

Security Requirements Language

CanReach(Src,

SrcProfile,

Dest,

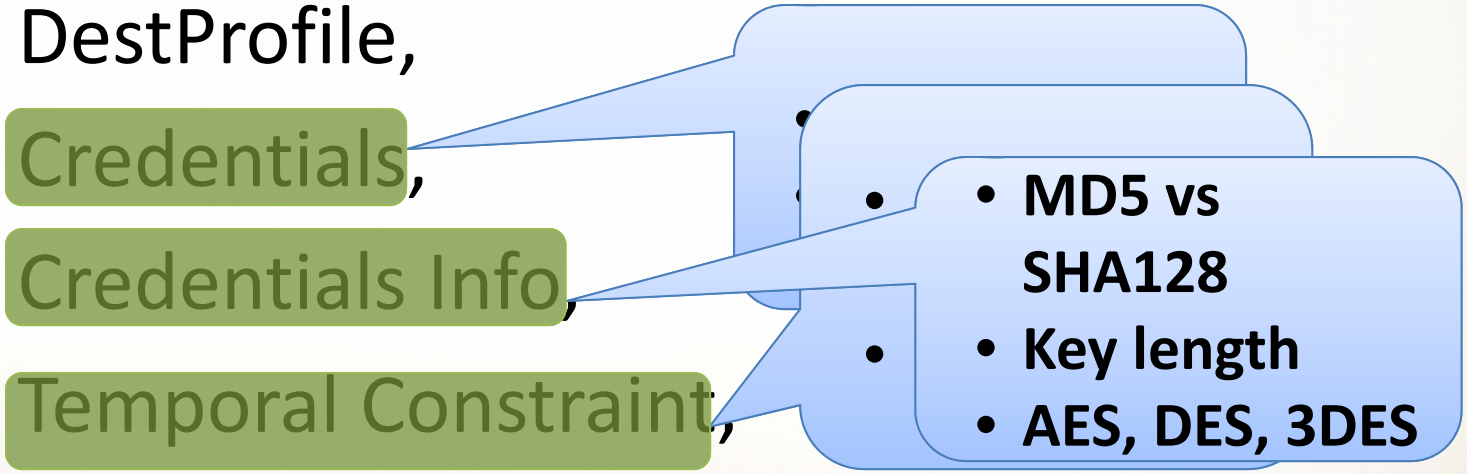
DestProfile,

Credentials,

Credentials Info,

Temporal Constraint,

Temporal Condition)



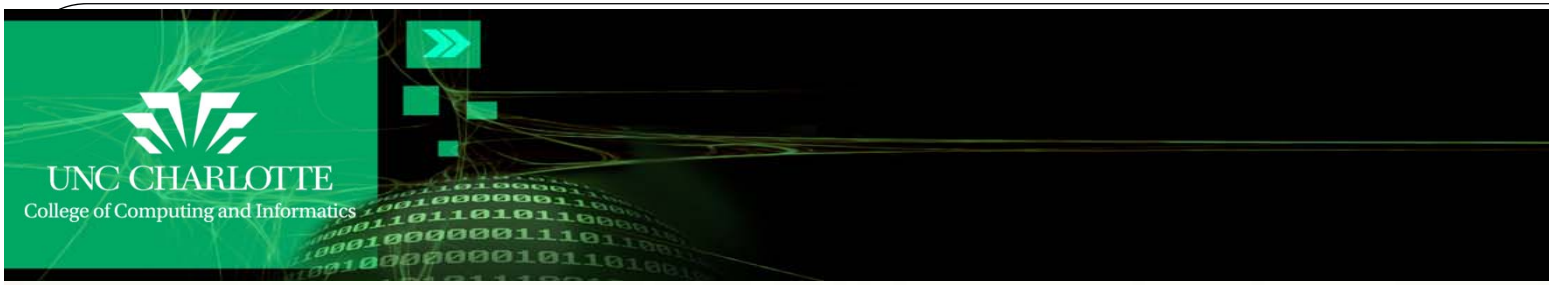
UNC CHARLOTTE
College of Computing and Informatics

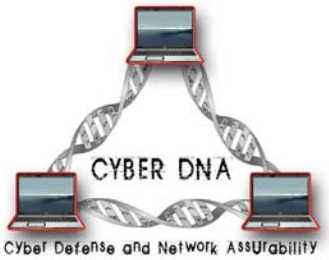
Security Requirements Language

- Logical Operators
  - AND
  - OR
  - NOT
  - IF-THEN (Implies).
- Temporal Operators
  - Always
  - Until
  - Unless

The traffic going to the SQL server should be inspected unless it is encrypted.

```
CanReach( Host,
          _,
          SQLServer,
          _,
          Inspected,
          _,
          Unless,
          Encrypted)
```





FireBlanket: On Synthesizing Distributed Filtering Configuration Considering Risk, Usability and Cost Constraints

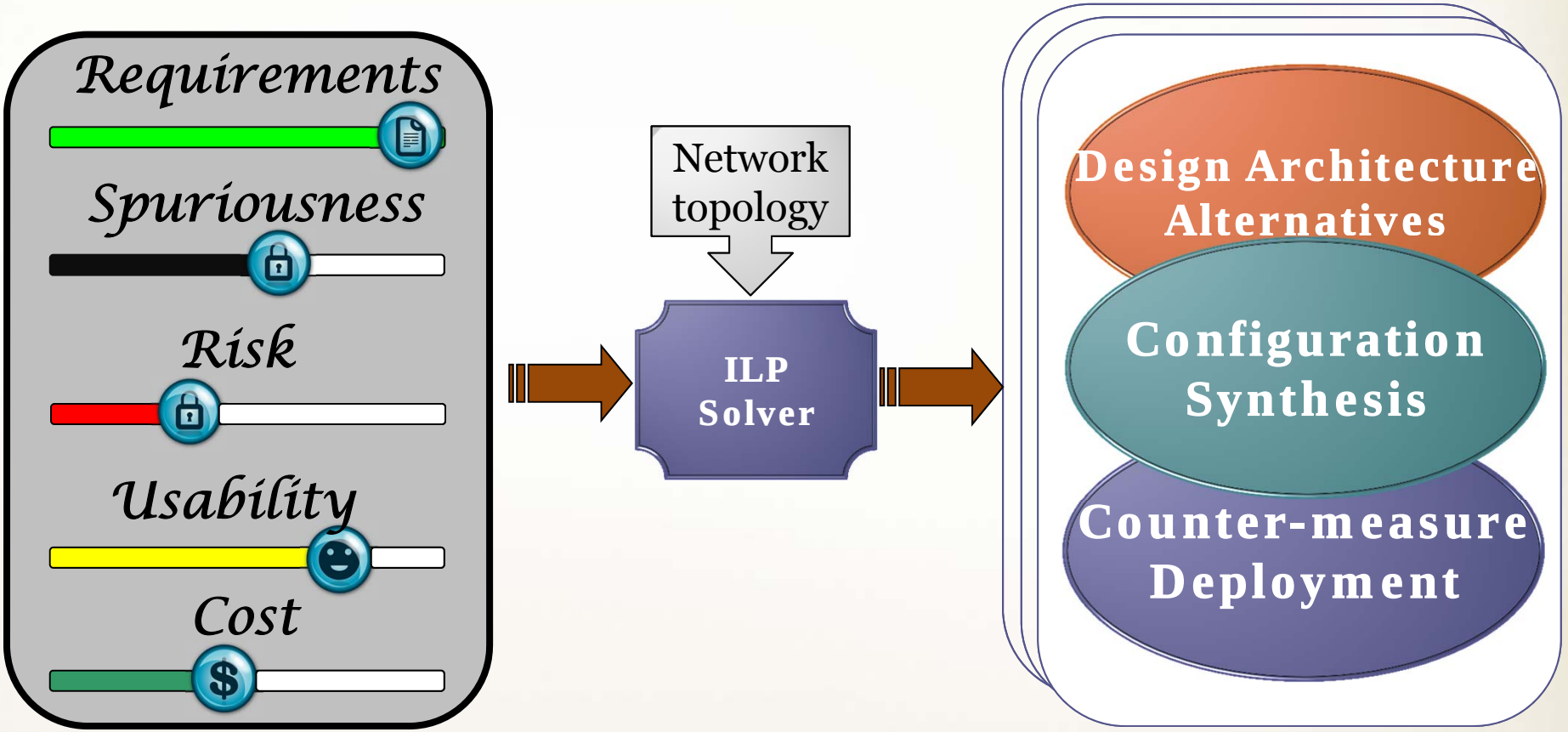
TOP DOWN Security Analysis

IEEE Conference on Network and Service Management (CNMS 2011)

UNC CHARLOTTE
College of Computing and Informatics

CONFIGSLIDER: The Science of Objective Security Configuration Synthesis

- ConfigSlider: Multi-factor network design tool ➔ what is the most cost-effective investment for maximizing security – design space exploration



UNC CHARLOTTE
College of Computing and Informatics

Research/Project Overview

Tool	Theory	Domain	Goal	Publication
FirewallPolicyAdvisor	Set theory	Firewalls	Consistency Checker	INFOCOM 04, IM03, JSAC05
ScuritylPolicyAdvisor	BDD	Firewall & IPSec	Consistency Checker	ICNP05
INSPEC	BDD	Firewall	Automated Testing for Cisco	JSAC09
FWPolicyVis	BDD	Firewall	Visualization	Usinex07
ConfigChecker	BDD-Model Checker	Global Enterprise	Automated Verification, diagnosis, and repair	ICNP09
ConfigLEGO	BDD & SAT	Global Enterprise & Clouds	Imperative verification, diagnosis, and repair	SecureComm 11 On progress
FireBlanket	ILP	Distributed Firewall	Automatic synthesis and DMZ creation	INFOCOM10 & CNSM11
ConfigSlider	SMT	Global	Automatic synthesis of global security architecture & config	On progress
CloudChecker	BDD/SMT	Cloud	Verification, compliance checking and Planning	On progress
SensorChecker	BDD-model Checker	WSN	Reachability and coverage verification	On progress
SensorPlanner	SMT	CPS	Coverage and mission satisfaction	On progress
ConfigSeal	SAT-Model Checker	Enterprise Logs and config	Verification and forensic to trackback	On progress

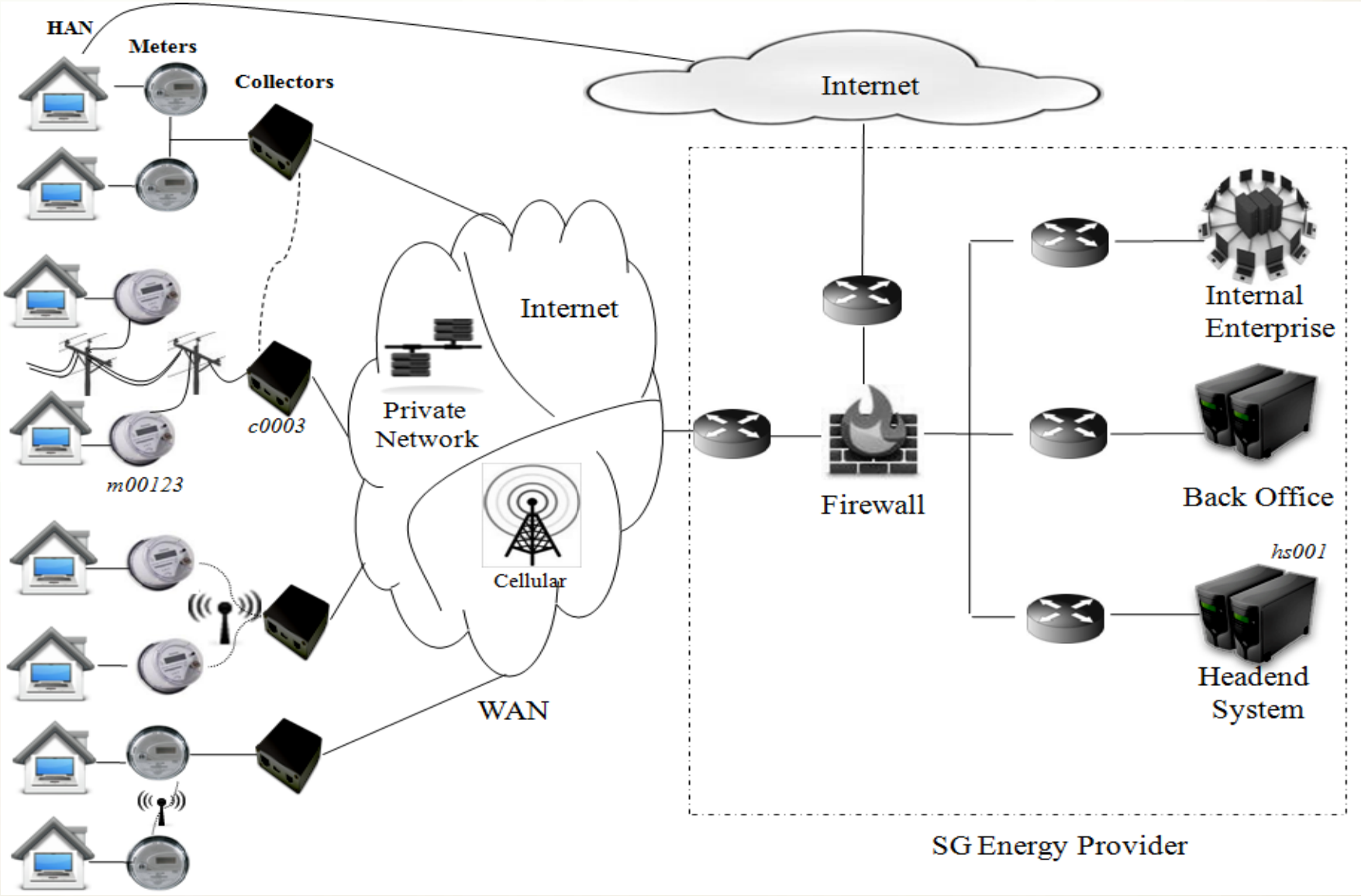
UNC CHARLOTTE
College of Computing and Informatics

SmartAnalyzer: A Noninvasive Security Threat Analyzer for AMI Smart Grid

IEEE INFOCOM'12, March 26-30, Orlando, FL

UNC CHARLOTTE
College of Computing and Informatics

Advanced Metering Infrastructure (AMI)



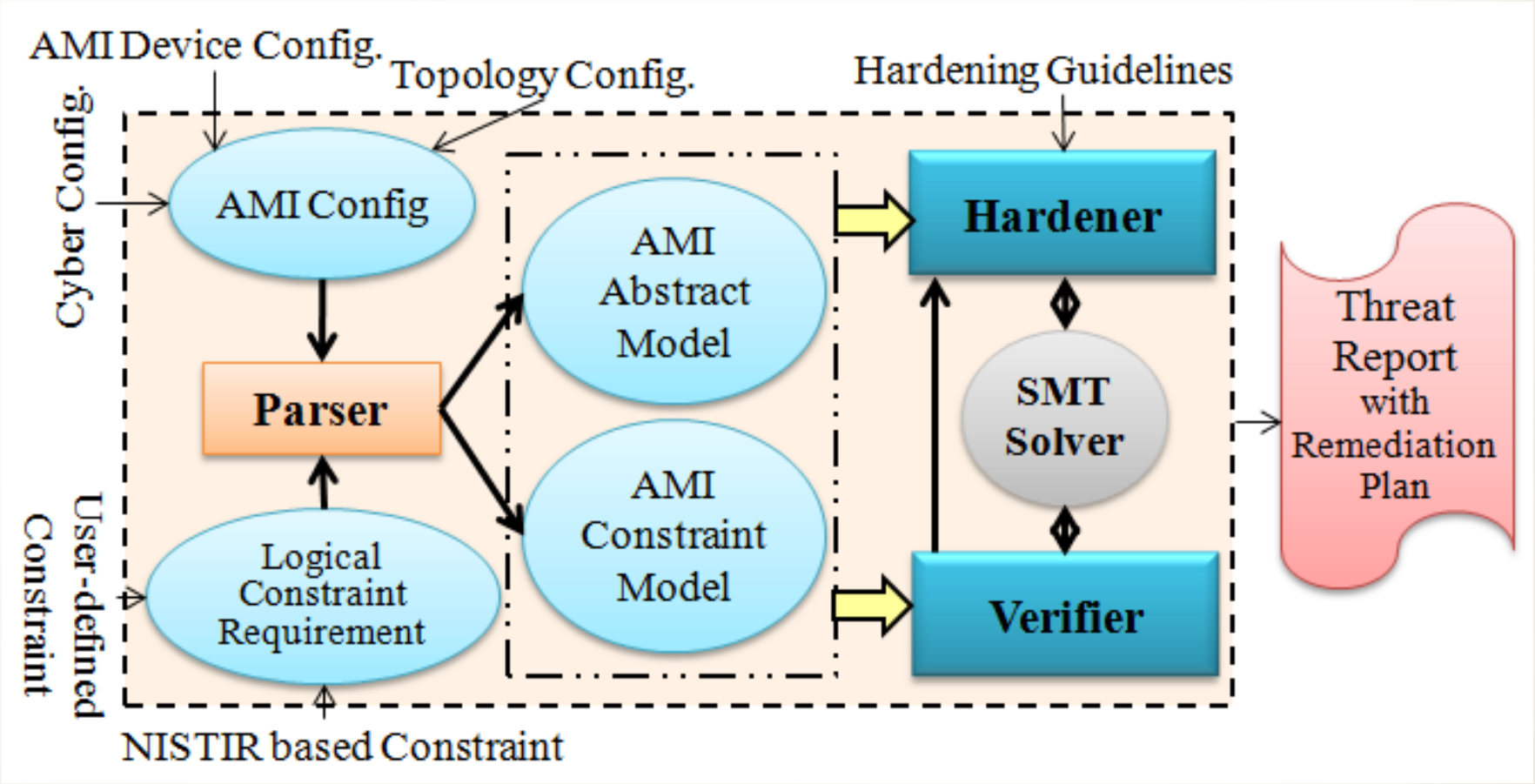
UNC CHARLOTTE
College of Computing and Informatics

Contributions

- Salable modeling of AMI configuration using SAT/SMT.
- Defining logic-based operational constraints for **AMI security and safety invariants** for smart grid. Examples:
  - Data overwrite protection
  - Device scheduling and cyber bandwidth constraint
  - Assured data delivery and data freshness
- Developing an automated AMI analytical tool based on SMT that allows to **objectively** (with proofs) assess and investigate AMI security configuration *for identifying and mitigating potential security threats proactively.*
- Provide proof-based threat and diagnosis reports of security violations or potential threats.

UNC CHARLOTTE
College of Computing and Informatics

Security Threat Analyzer for AMI Smart Grid



UNC CHARLOTTE
College of Computing and Informatics

Challenges & Suggested Projects

- **Cloud compliance checking (from both user and provider prospective)**
- **Characterization of Attack/Failures Root Cause on Mideast ISP**
  - **Lack of protection**
  - **Misconfiguration**
  - **Recoverable Vulnerabilities**
  - **0-day attacks**
- **Global Consistency (firewall configuration consistency ==> ISP/University anonymous firewall configuration**
- **formal analytics of network/system logs for security compliance ==> ISP/University logs**
- **misconfiguration identification using traffic analysis ==> requires ISP traffic**
- **Data = Black Gold**

UNC CHARLOTTE
College of Computing and Informatics

Food for Thoughts

- **NSF SafeConfig Workshop Report, 2009**
  - **www.safeconfig.org**
- **Security Automation Research: Challenges and Future Directions**, DoD Information Assurance Newsletter, Information Assurance Technology Analysis Center, Volume: 14, Number: 4, Pages: 14-18, December 2011