NetPrints

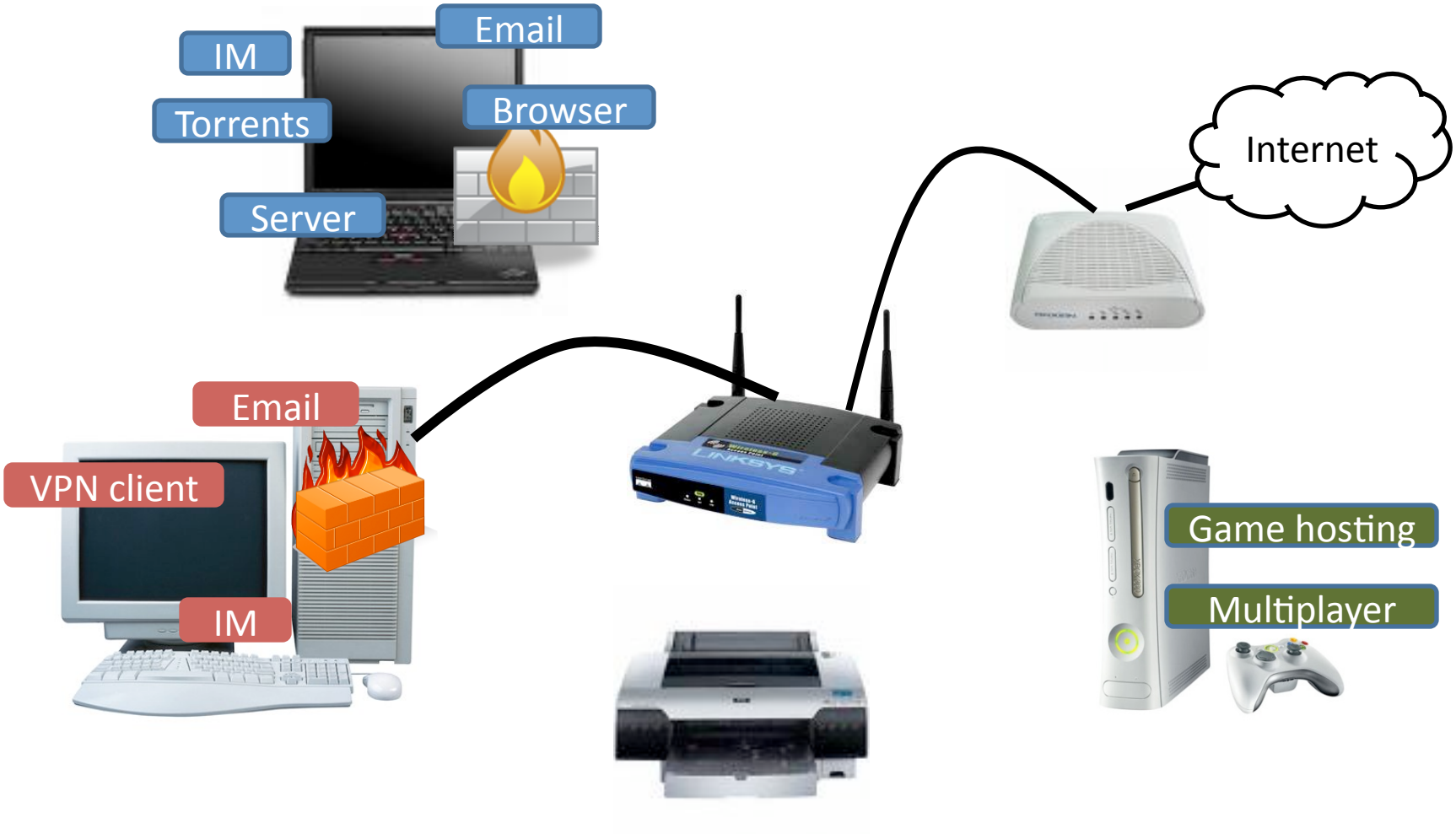
Diagnosing Home Network Misconfigurations using Shared Knowledge

Bhavish Aggarwal, Ranjita Bhagwan, Tathagata Das, Venkat Padmanabhan
Microsoft Research India

Siddharth Eswaran, **IIT Delhi**
Geoff Voelker, **UCSD**

Microsoft Research india

Typical Home Network



No network admin!

Examples of Problems

Problem	Solution
VPN client does not connect from home	Turn on PPTP passthrough on router, use a subnet that is either 192.168.0.x or 192.168.1.x
XBOX doesn't connect to the Live service	Turn up your MTU above 1365, set NAT settings to full-cone, turn on UPnP
My IM client doesn't work from home	Turn off the DNS proxy on the router
File sharing doesn't seem to work at home	Make sure you and the file server are on the same domain/workgroup
Printing doesn't work from my laptop	Turn on correct firewall rules on the server machine
Cannot send large emails	Turn down MTU on your router

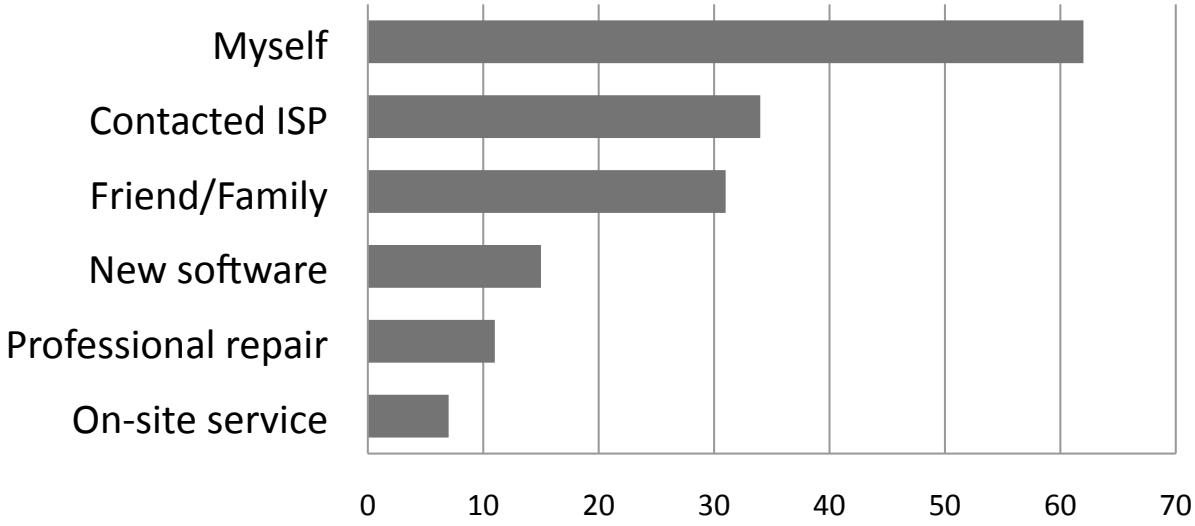
Router misconfig

End-host misconfig

Remote problem, local changes

Diversity ⇒ home network troubleshooting is hard

What Do Users Do Today?



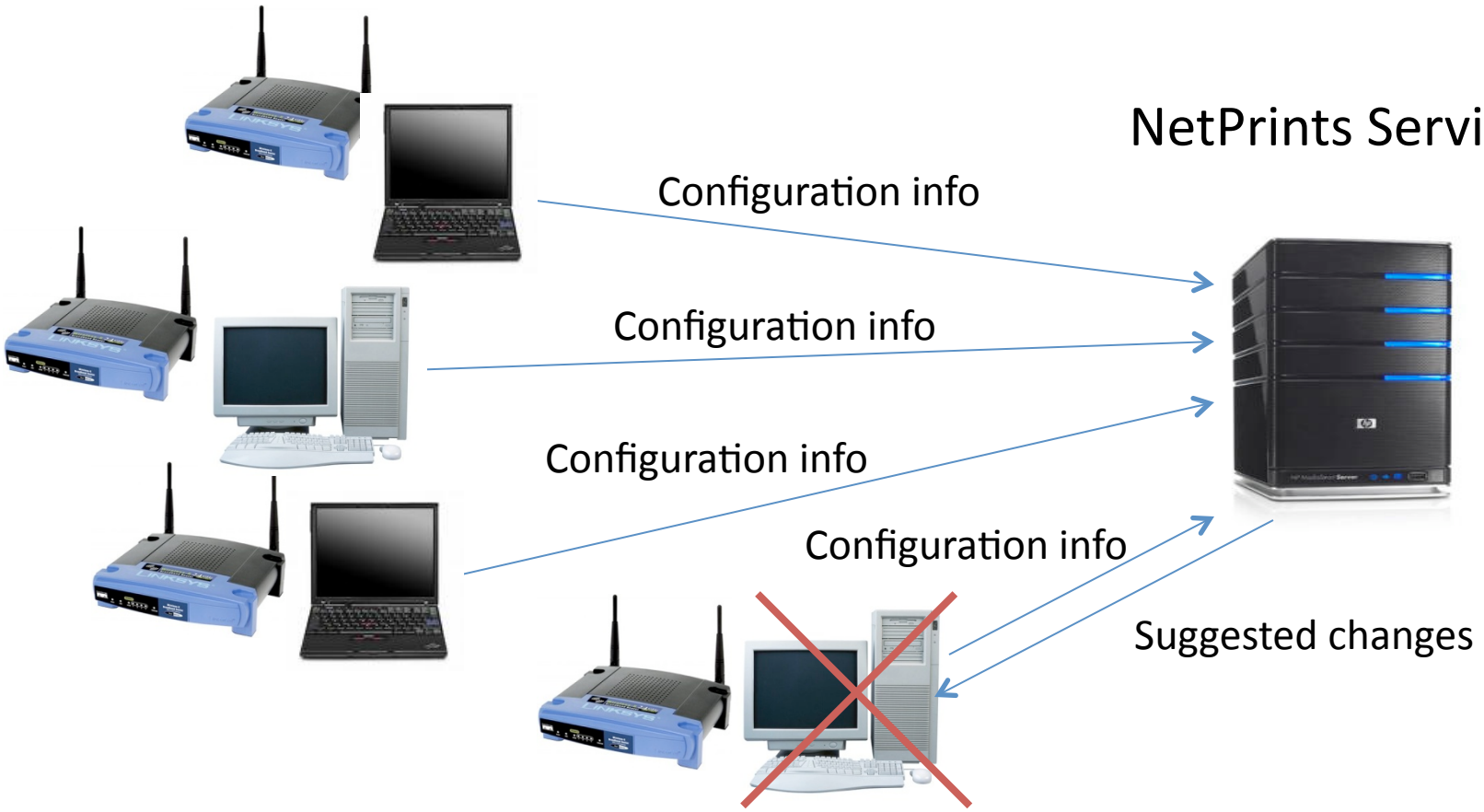
Source: Managing the Digital Home, a survey of 6,116 U.S. and Canadian home Internet users

Avg time to resolve solutions: 2 hours

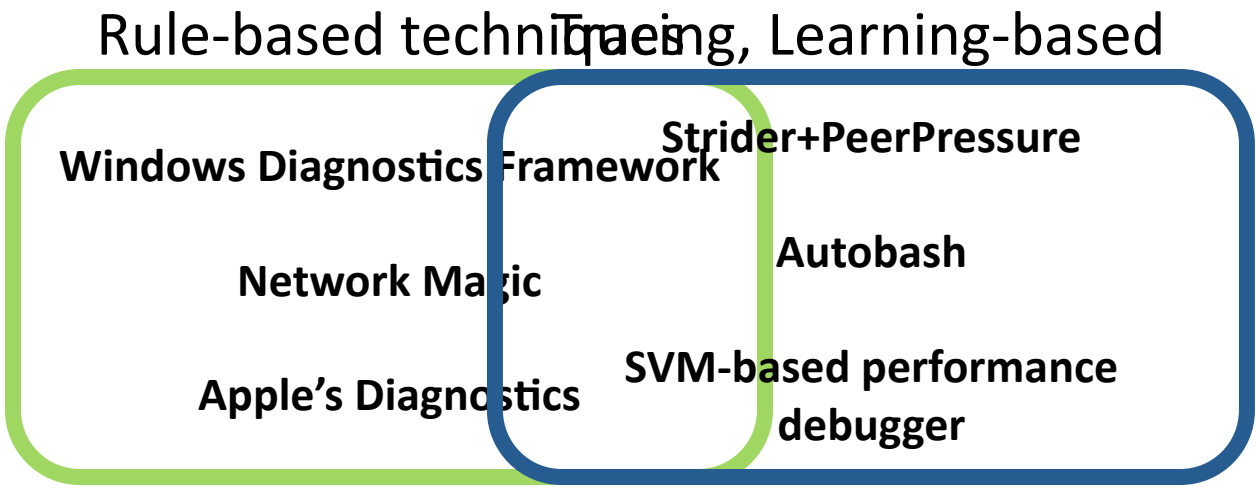
NetPrints

NetPrints = Network Problem Fingerprinting

Automate problem diagnosis using "*shared knowledge*"



Putting NetPrints in Context



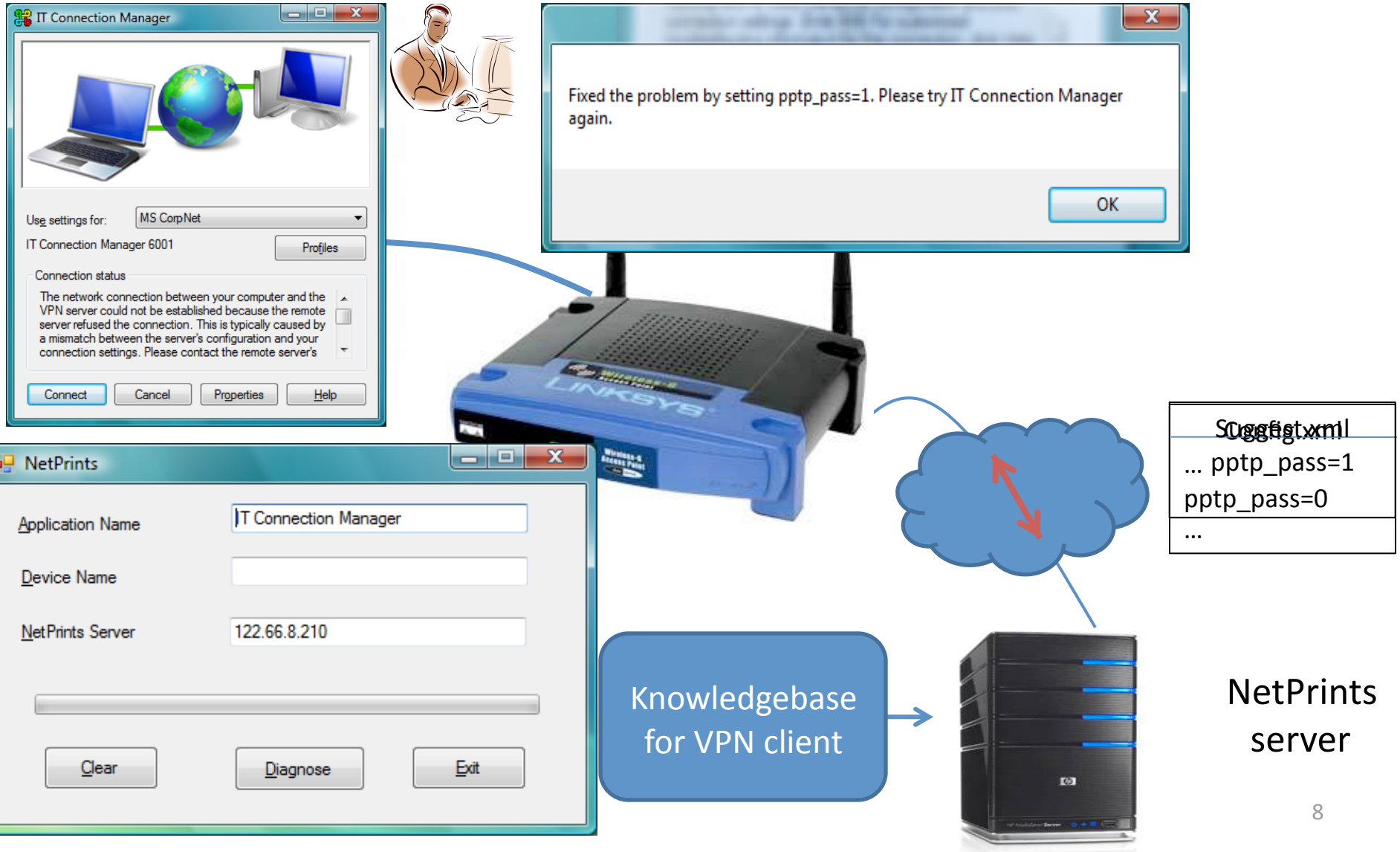
NetPrints

- Distributed configuration information
- Unstructured, heterogeneous environment
- Problems caused due to interaction of multiple configurations

Assumptions

- Current design requires basic connectivity
  - Looking at application-specific problems
  - Not inherent, Knowledgebase can be shipped offline
- Not dealing with performance
  - “good” and “bad” are the only two states considered

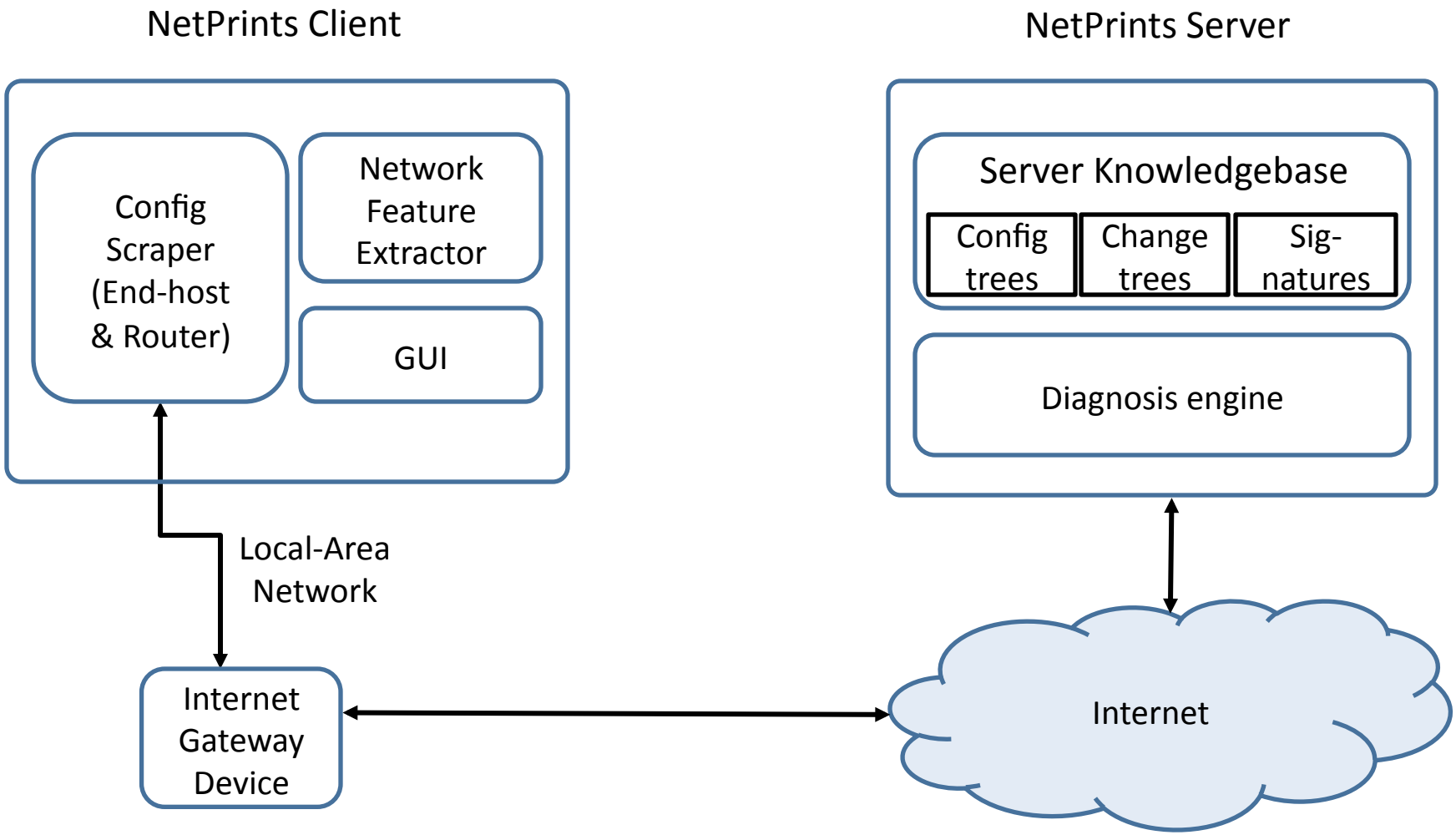
NetPrints in Action



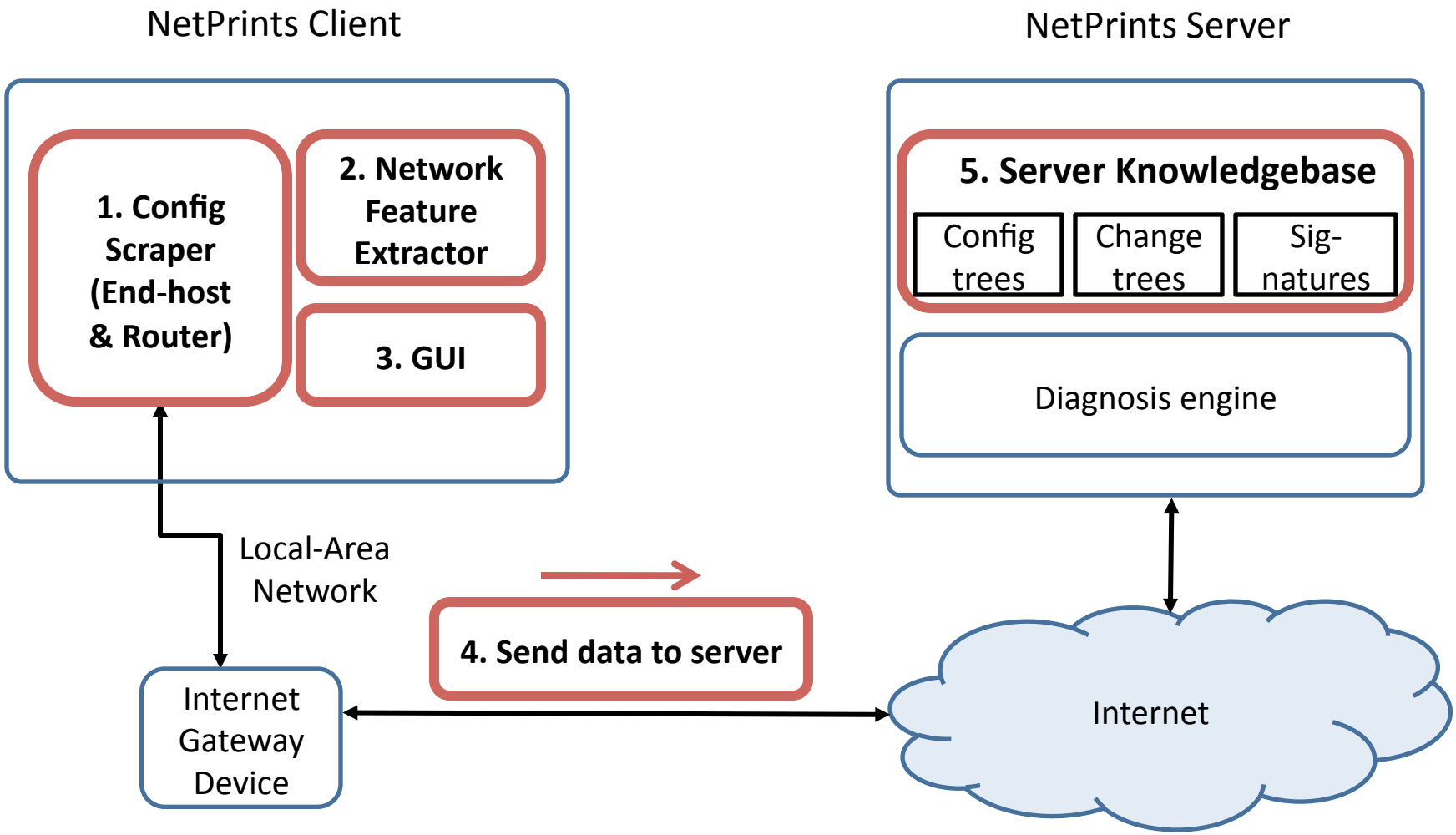
Diagnosis Strategies

- Snapshot-based
  - Collect config snapshots from different users
- Change-based
  - Collect config changes that a user makes
- Symptom-based
  - Collect signatures of problems from network traffic

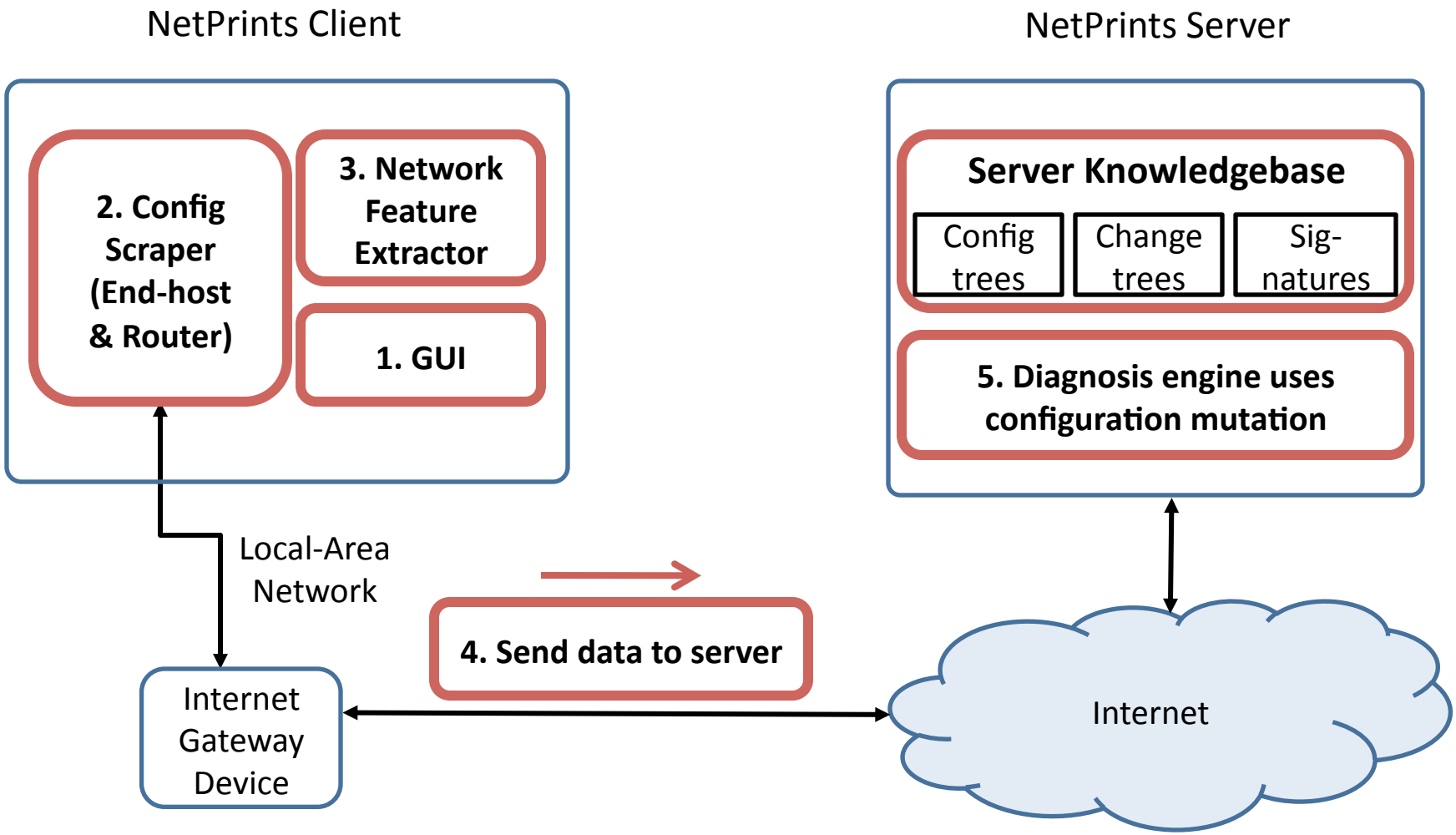
System Design



Normal Mode


NetPrints Client
NetPrints Server
1. Config Scraper (End-host & Router)
2. Network Feature Extractor
3. GUI
5. Server Knowledgebase
Config trees
Change trees
Sig-natures
Diagnosis engine
Local-Area Network
4. Send data to server
Internet Gateway Device
Internet

Diagnose Mode



#1: Configuration Scraper

- Router scraper
  - UPnP
  - Web Interface (HTTP Request Hijacking)
- End-host scraper
  - Interface-specific parameters
  - Patches and software versions
  - Firewall rules
- Remote scraper
  - Composition of local and remote configs

Composing Local & Remote Configs

Problem	Solution
VPN client does not connect from home	Turn on PPTP passthrough on router, use a subnet that is either 192.168.0.x or 192.168.1.x
XBOX doesn't connect to the Live service	Turn up your MTU above 1365, change NAT settings to full-cone, turn on UPnP
My IM client doesn't work from home	Turn off the DNS proxy on the router
File sharing doesn't seem to work at home	Make sure client and the server are on the same domain/workgroup.
Printing doesn't work from my laptop	Turn on correct firewall rules on print server machine
Cannot send large emails	Turn down MTU on your router

Sometimes it is the *combination* of local and remote configs that is the problem

#2: Server Knowledgebase

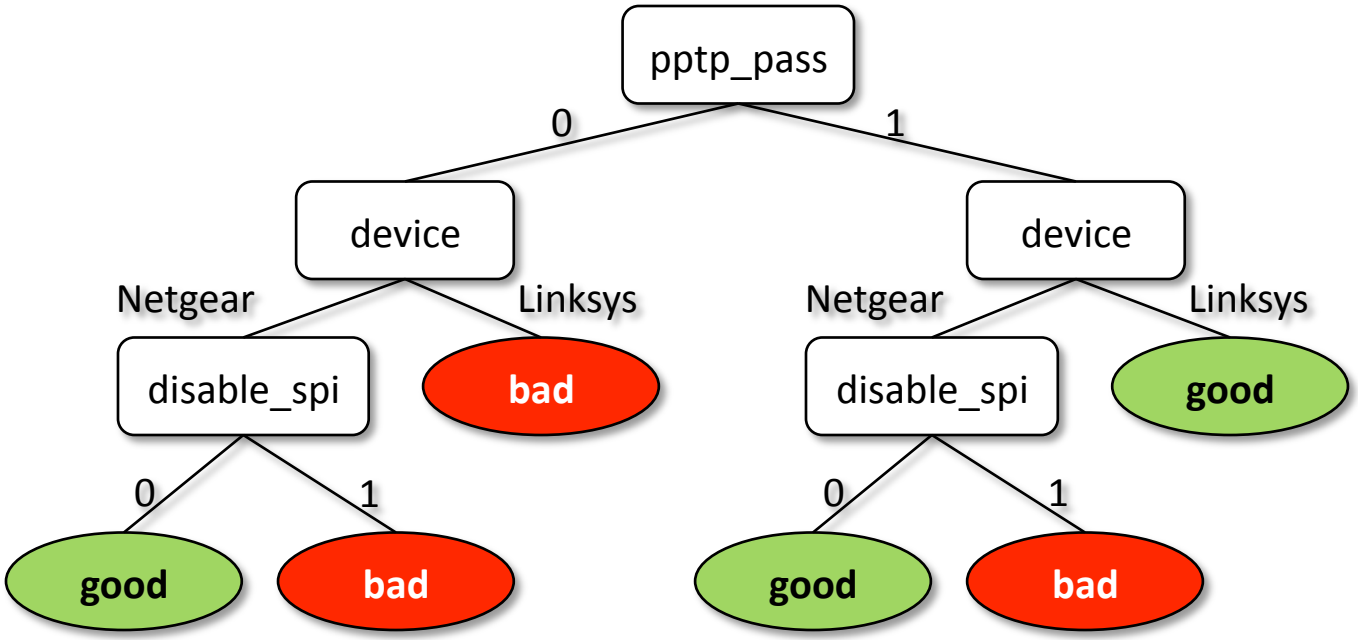
- Per-application decision trees constructed using labeled configuration snapshots
  - decision trees aid interpretability
  - C4.5 decision tree learning algorithm

- Configuration tree, Change trees and network signatures

Methodology

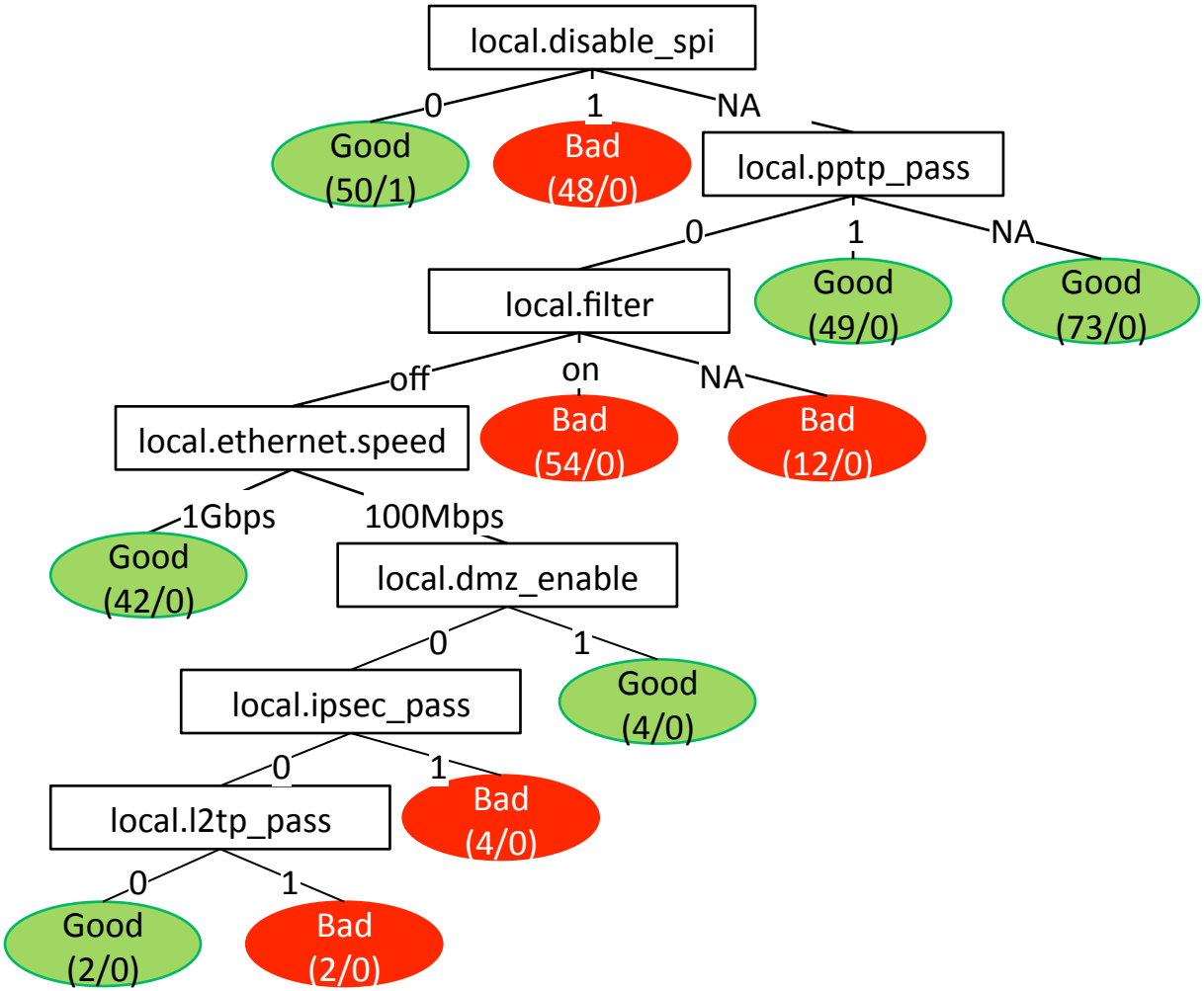
- Testbed comprising 7 different routers
  - various makes: Netgear, Linksys, D-Link, Belkin
- Clients running the VPN sent configurations to the NetPrints service
  - Roughly 6000 config parameters per snapshot
- Service learned configuration trees using C4.5 algorithm

Example of Configuration Tree



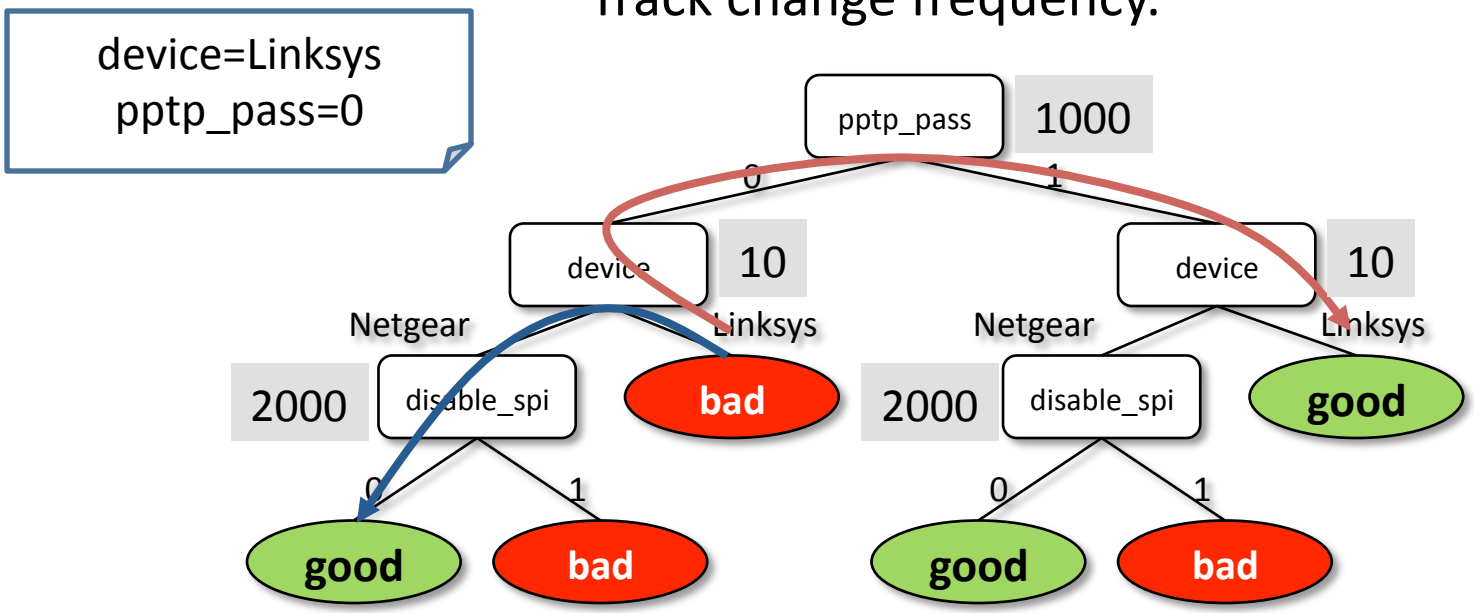
Simplified Config Tree for VPN Client

Configuration Tree for VPN Client



#3: Configuration Mutation

Track change frequency.

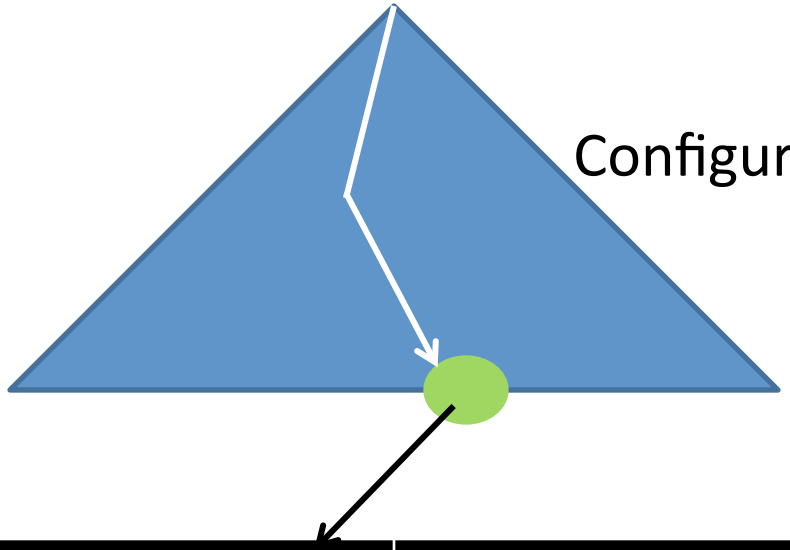


- Preference for mutations involving frequently changing parameters
- Assumption: higher the frequency, less disruptive the change

Shortcoming of Configuration Trees

- Some config info may not be learned
- So traversal of config tree may end in a "good" leaf even if config is problematic
- Reasons:
  - **Insufficient data**
    - e.g., a new router enters the market
  - **Hidden configurations**
    - e.g., application-specific parameters

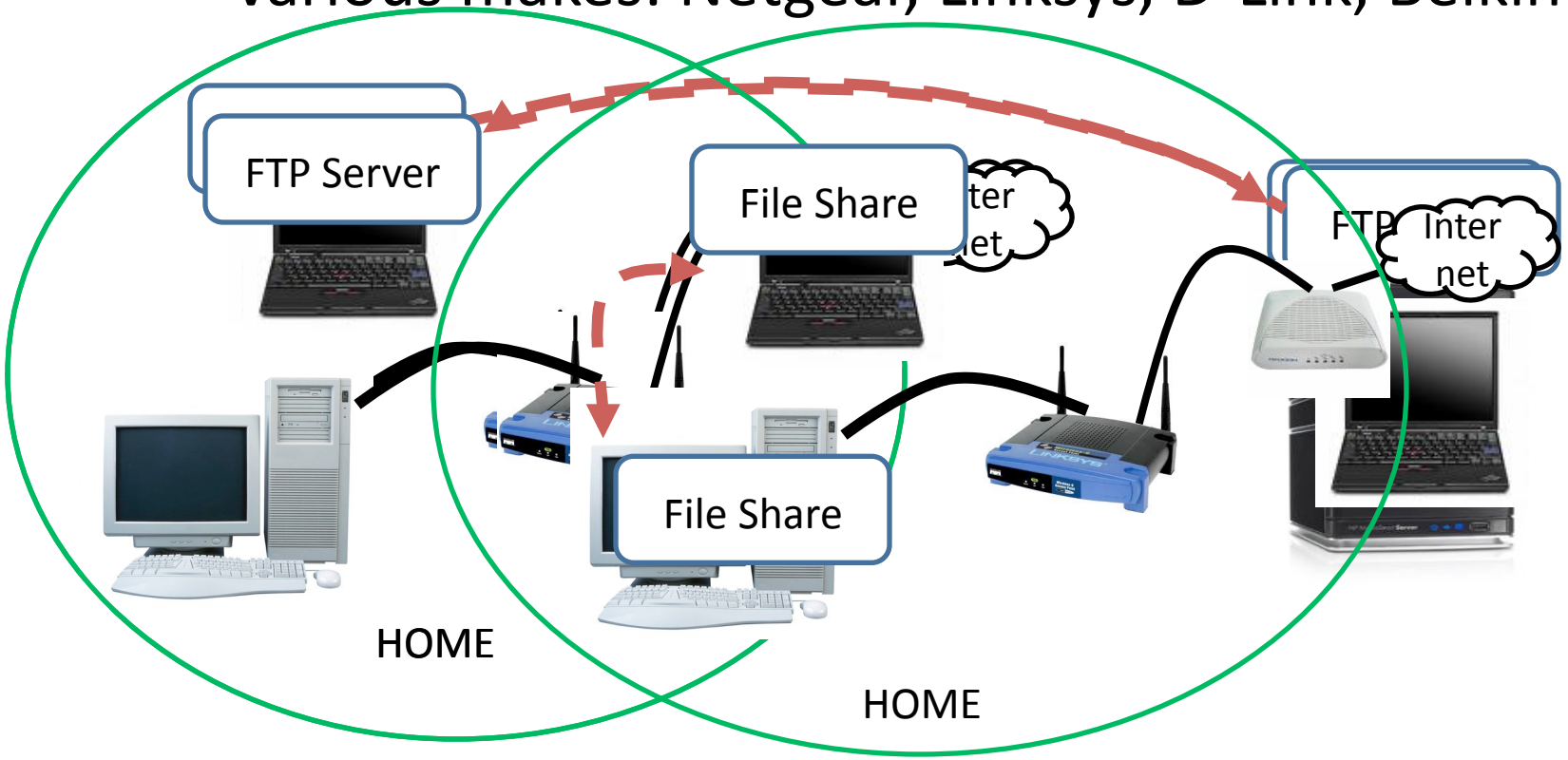
Summary of Diagnosis Procedure



Network traffic signature	Change trees
1 X X X X X X	
0 X X X X 1 X	

Experimental Evaluation

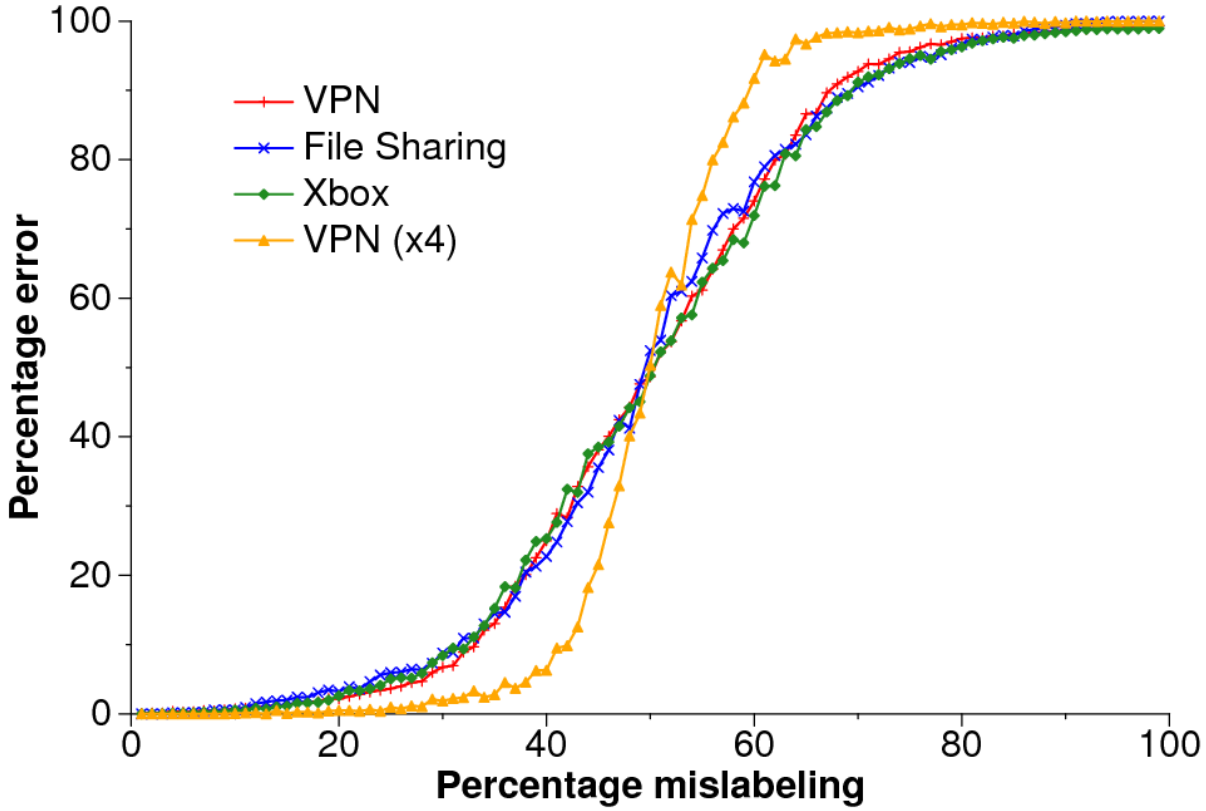
- Testbed comprising 7 different routers
  - various makes: Netgear, Linksys, D-Link, Belkin



Findings

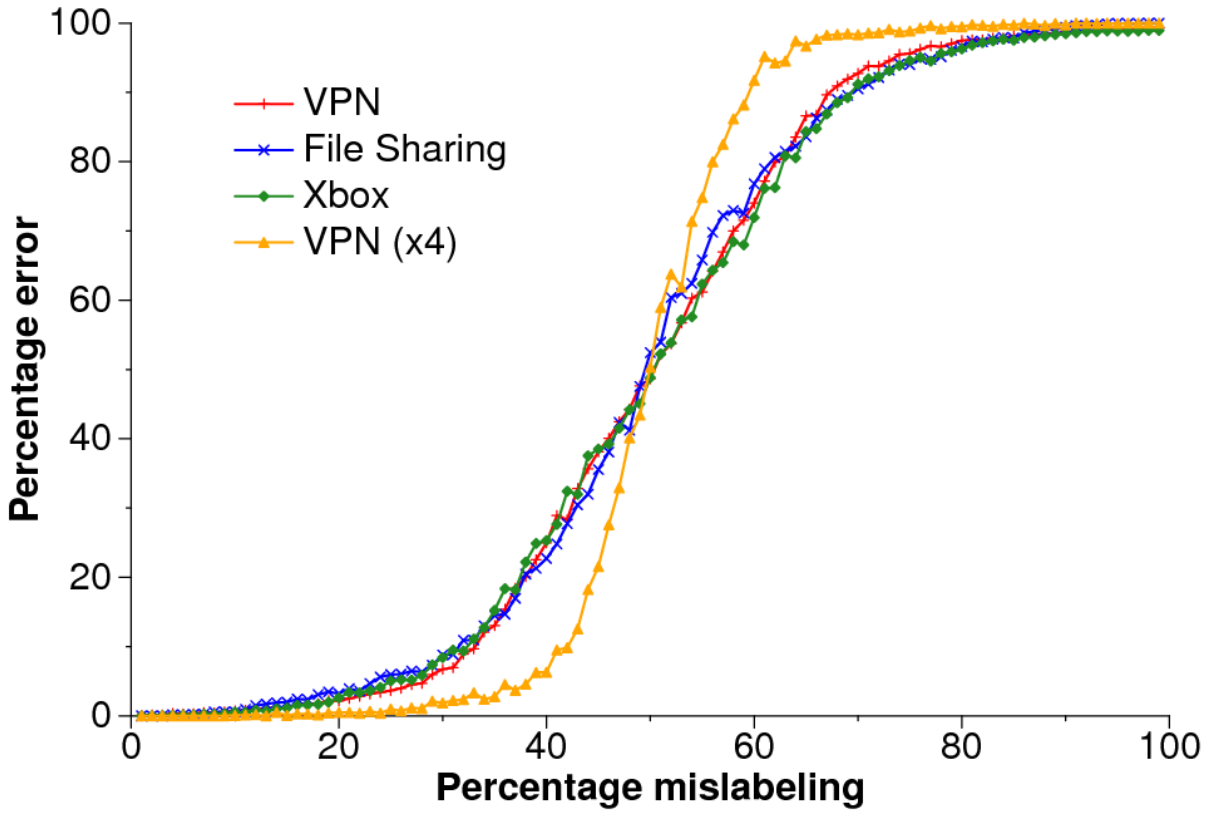
- Intuitive inferences
  - VPN: If **pptp_pass==1** then GOOD
- Surprising inferences
  - VPN: If **stateful==off** and **pptp_pass==0** and **ipsec_pass==0** and **l2tp_pass==0** then GOOD

Tolerance to Mislabeling



13-17% mislabeling ⇒ 1% error in diagnosis

Tolerance to Mislabeling



13-17% mislabeling ⇒ 1% error in diagnosis

Summary

- Home network diagnostics is challenging
  - diversity of apps and configs
  - absence of an admin
- NetPrints leverages community info to perform *automated* diagnosis
  - decision tree based learning
  - configuration trees, network traffic signatures and change trees

Thank you

bhagwan@microsoft.com

http://research.microsoft.com/netprints