

# Secure Remote Diagnostics

---

Justin Brickell

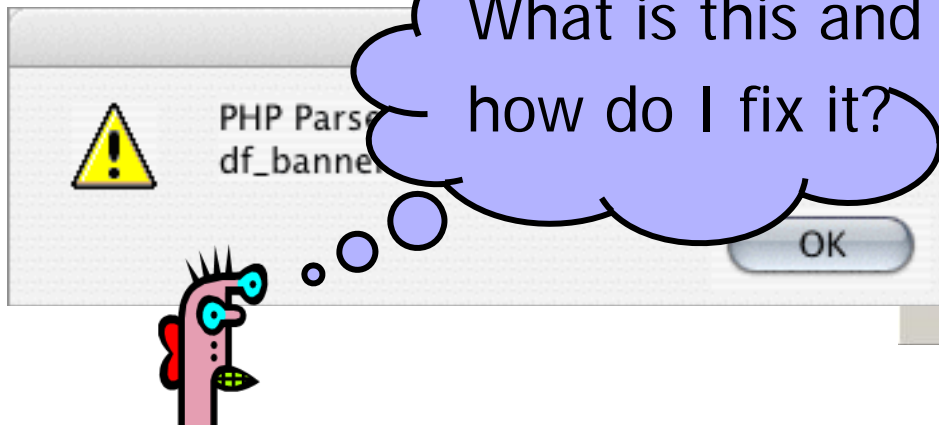
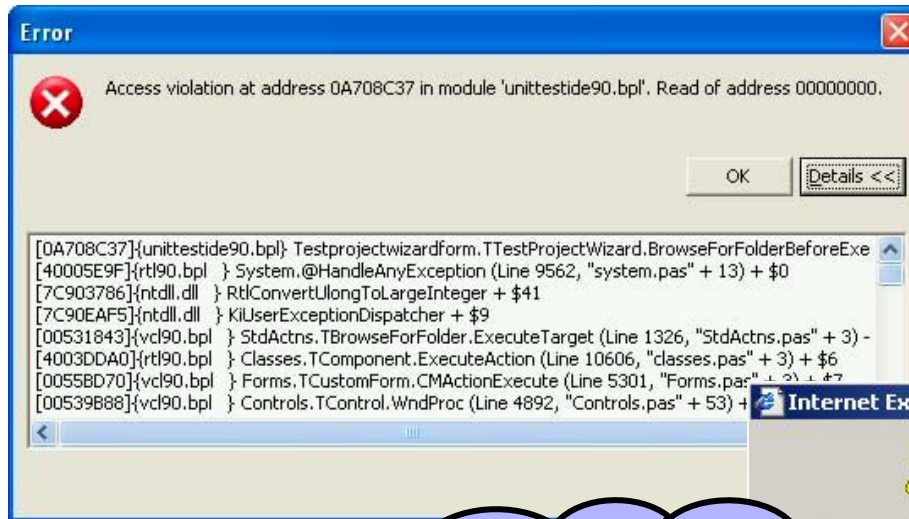
Donald E. Porter

Vitaly Shmatikov

Emmett Witchel

The University of Texas at Austin

# Error messages are cryptic



# Troubleshooting is no longer local

## DrWatson Postmortem Debugger

DrWatson Postmortem Debugger has encountered a problem and needs to close. We are sorry for the inconvenience.

If you were in the middle of something, the information you were working on might be lost.

### Please tell Microsoft about this problem.

We have created an error report that you can send to help us improve DrWatson Postmortem Debugger. We will treat this report as confidential and anonymous.

To see what data this error report contains, [click here](#).

I can diagnose this fault online



## Mozilla Quality Feedback Agent - Firefox10



The Mozilla Quality Feedback Agent has captured information that Mozilla needs to help improve Firefox's quality.

Enter your email address, describe how you were using Firefox, then click Send.

Your Email Address (optional):

Send me information about updates to Mozilla products

If failure occurred within the browser please provide URL/location:

Describe what you were doing when Firefox failed (optional):

A nagytalpú bűdöskurvaúristen megbasszaazegét!

Automatically send incidents (don't show this dialog again)

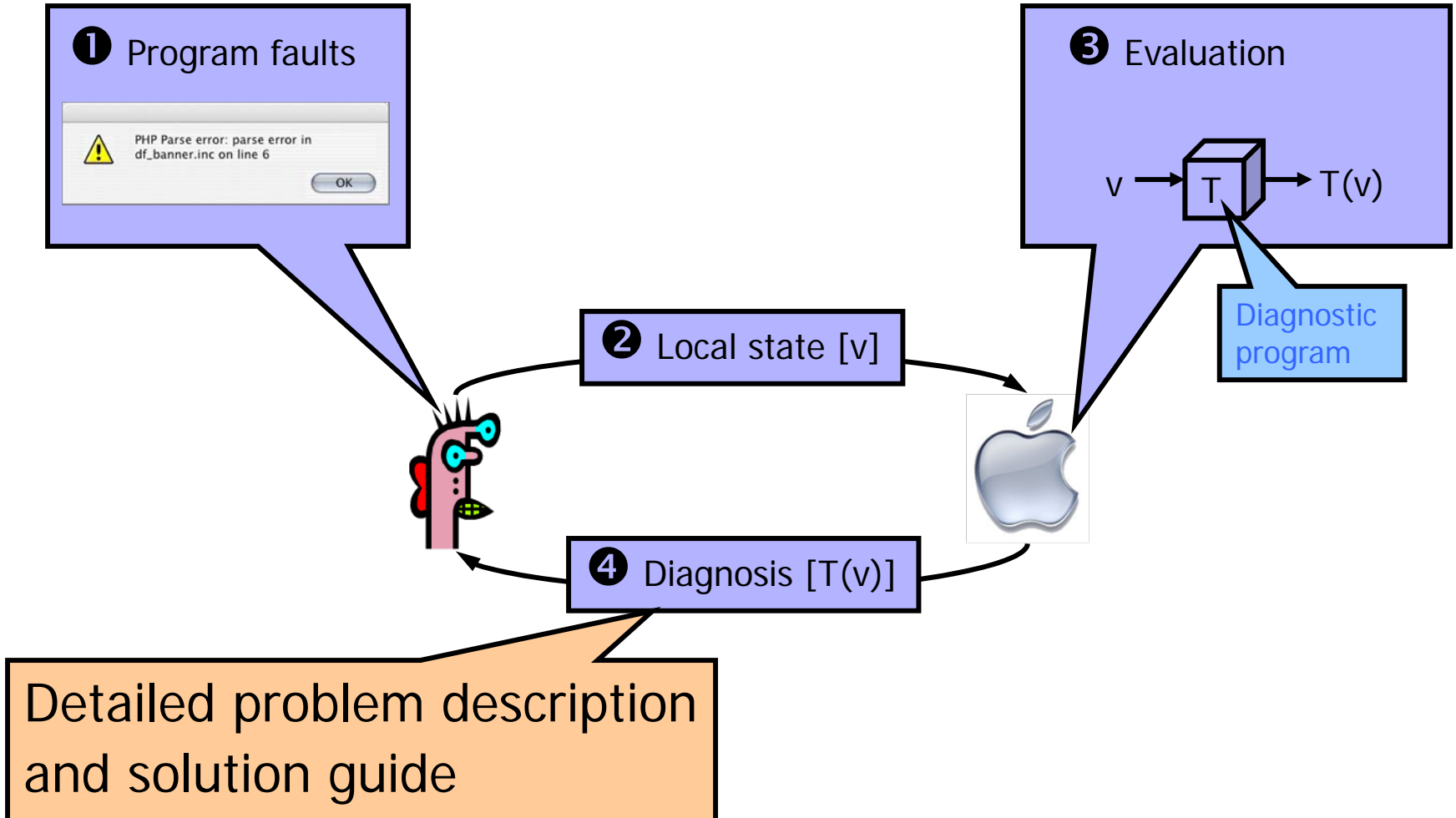
The Mozilla Quality Feedback Agent is based on SupportSoft Talkback is a Trademark of [SupportSoft Software, Inc.](#)

Send

Don't Send

Show Details

# Software diagnostic scenario



# Our goal: Protect user privacy

- Program memory
- File snippets
- Other running programs
- Confidential information



“If you are concerned that a report might contain personal or confidential information, you should not send the report.”

Local state [v]



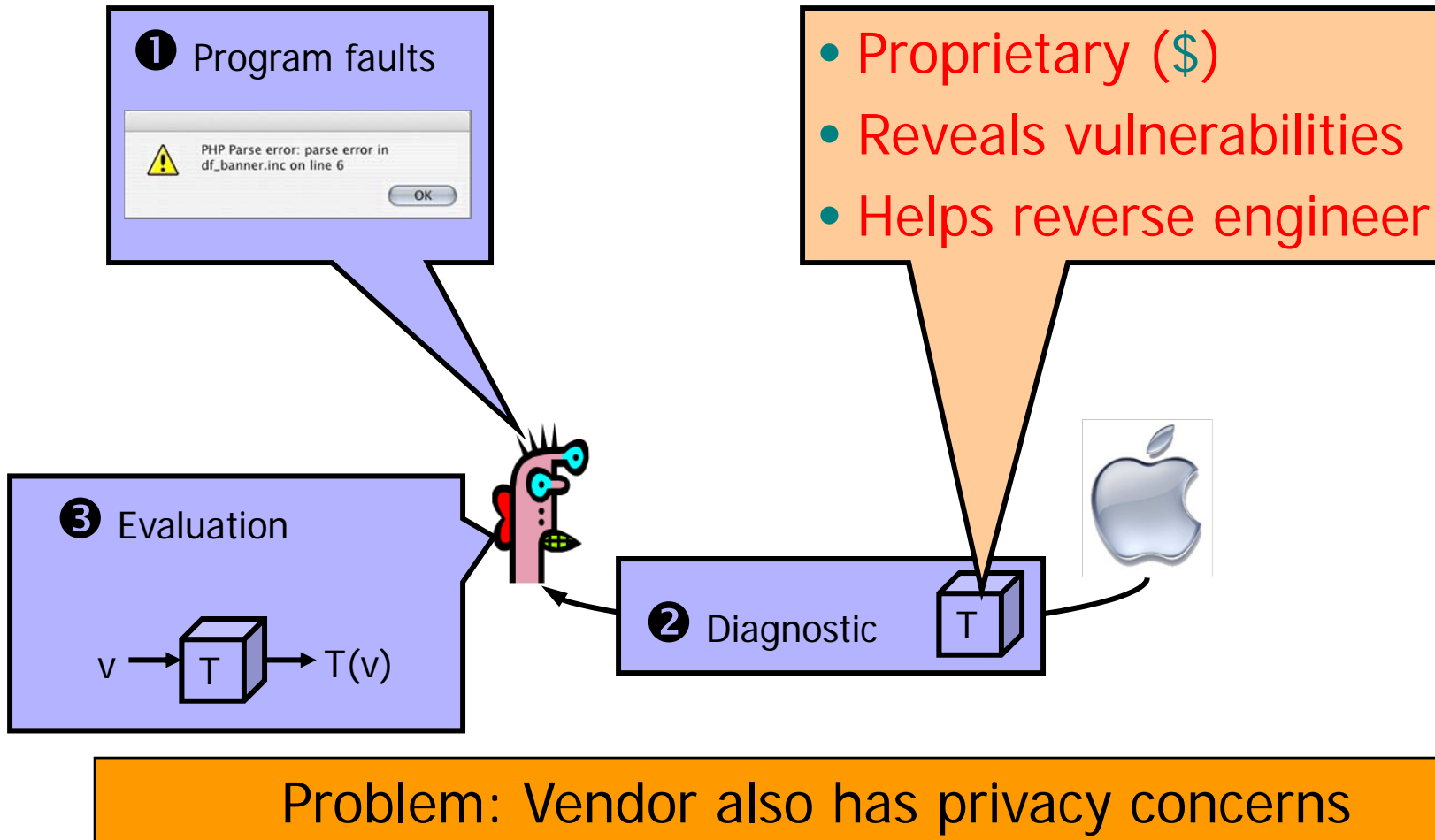
Problem: Users with privacy concerns cannot participate

# Talk outline

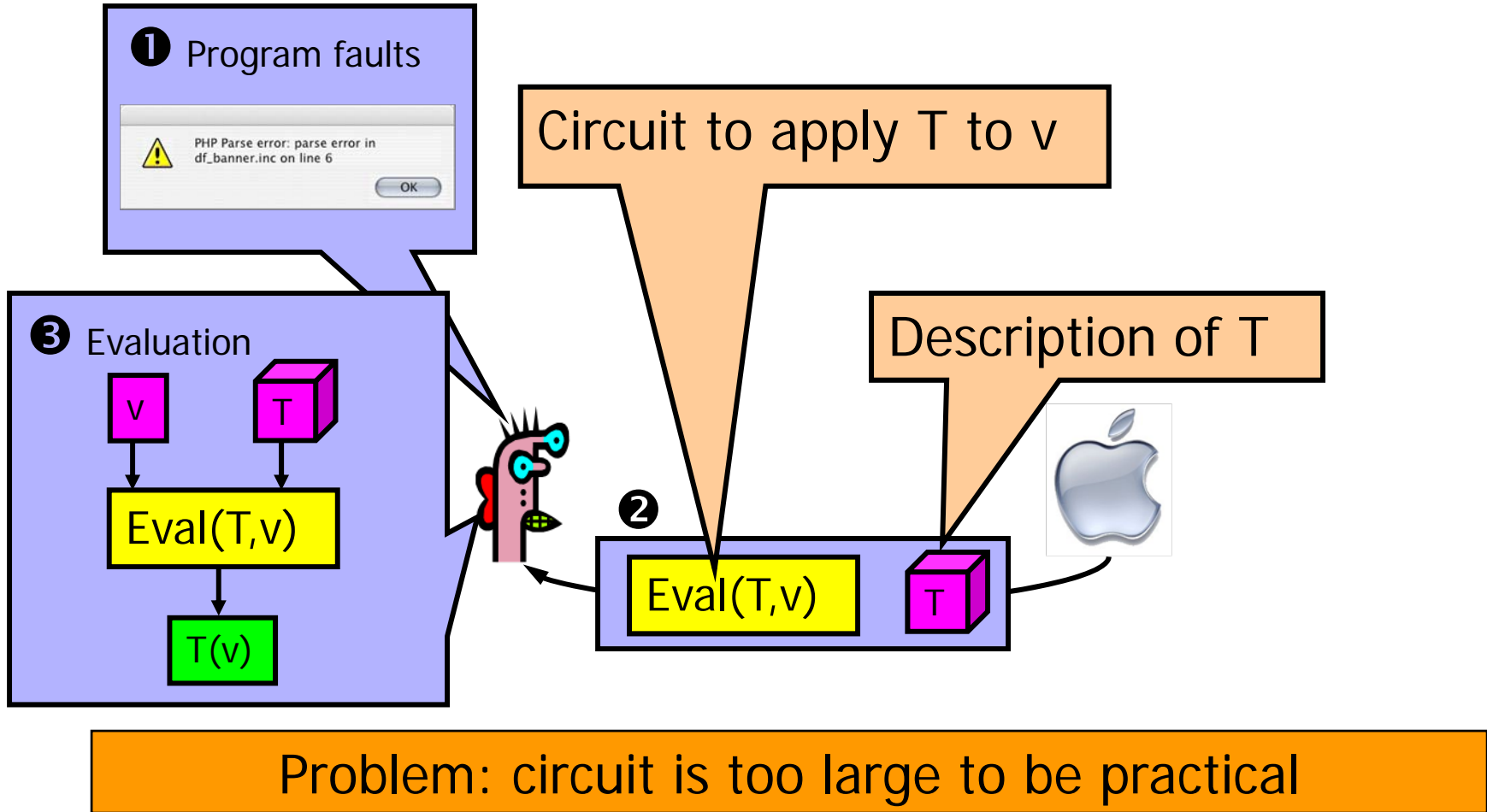
---

- ◆ Two unsatisfactory approaches
- ◆ Overview of our approach
- ◆ Building blocks
- ◆ Protocol walkthrough
- ◆ Applications
- ◆ Performance
- ◆ Conclusion

# User-side evaluation

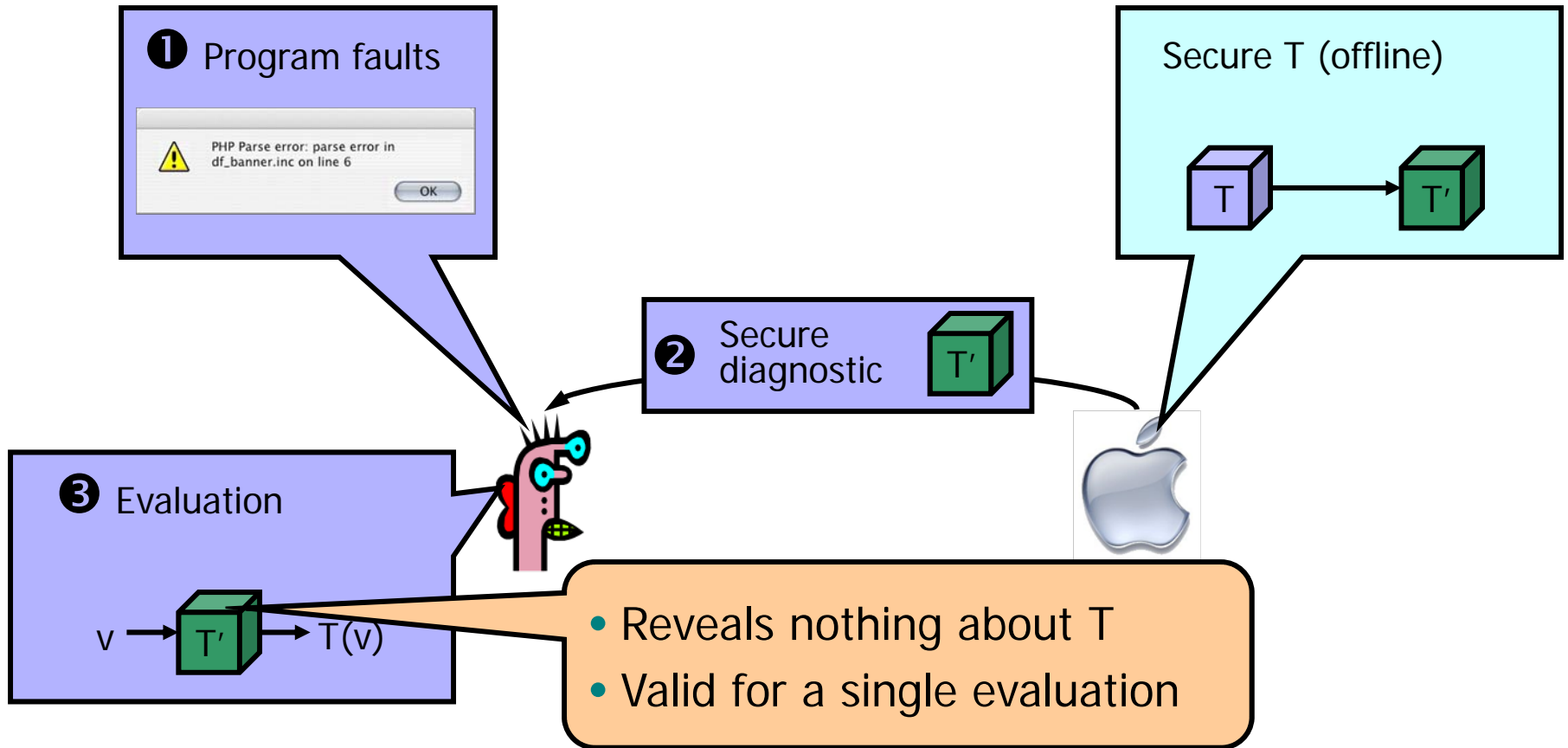


# Secure multiparty computation





# Our approach: Securing T



Goal: Build a **practical** system for real problems

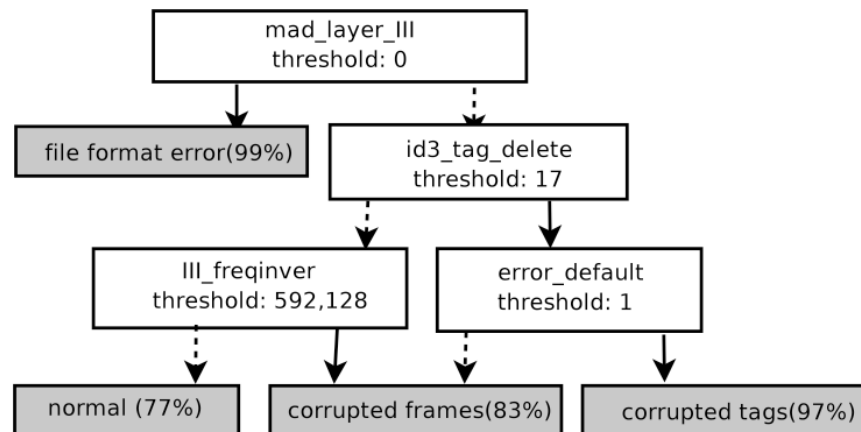
# Local state

---

- ◆ A snapshot of the user's local state
- ◆ A vector of attribute values, such as
  - Function call counts
  - Function return values
  - Contents of memory locations
  - ...

# Diagnostic branching program

- ◆ **Decision nodes** compare an attribute value to a threshold
  - Control flows to either the **left** or **right** branch
- ◆ **Classification nodes** specify a label



Diagnostic branching program for mpg321

# Aside: Creation of diagnostics

---

- ◆ Diagnostic programs are built in several ways
  - Hand-designed using expert knowledge
  - Automatically generated by data mining and analyzing user error reports
    - For example, Microsoft's Dr. Watson
- ◆ Clarify [PLDI '07]
  - Project at UT Austin to automatically build “black-box” classifiers for anomalous program behavior

# Hiding the diagnostic program

Property to hide	Technique
Per node: thresholds and attributes	Private integer comparison (Yao's method)
Global: subset of attributes that are used	Homomorphic encryption
	Blinding
Global: program topology	Unevaluated nodes are hidden by encryption

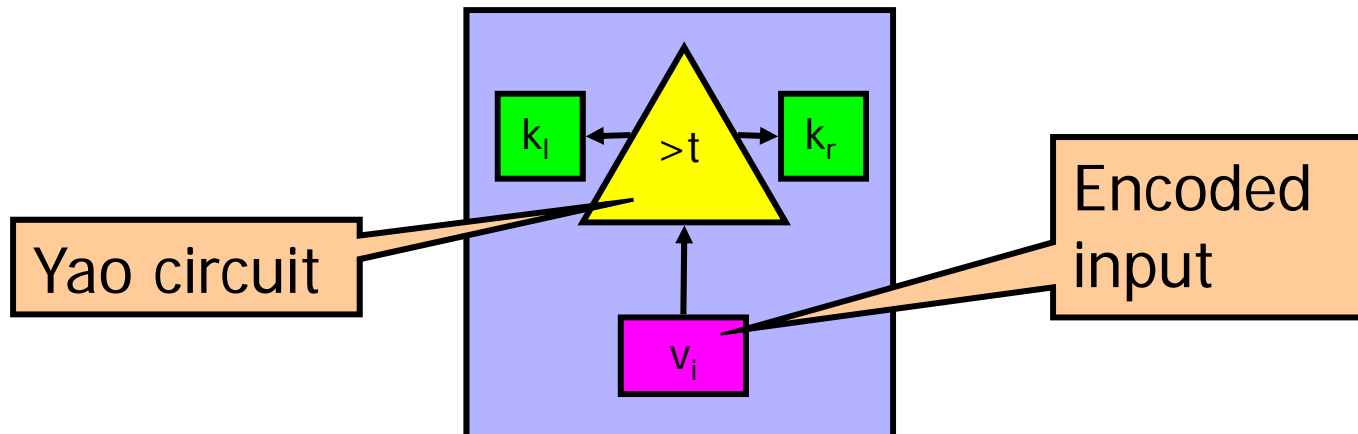
Hide everything about T

# Private integer comparison

◆ We use Yao's method for integer comparison

IF ( $v_i > \text{threshold}$ ) THEN  $k_l$  ELSE  $k_r$

- Evaluator learns one of two keys conditional on comparison result
- Evaluator doesn't learn comparison result or the threshold



# Protocol walkthrough

Hiding subset of attributes that are used

1. User encrypts attribute values using an instance of additively homomorphic encryption

Allows addition under encryption

$V_1$

$V_2$

$V_3$

$V_4$

$V_5$



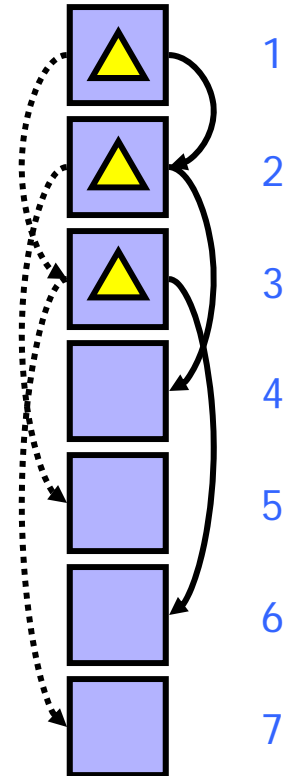
$V_1$

$V_2$

$V_3$

$V_4$

$V_5$

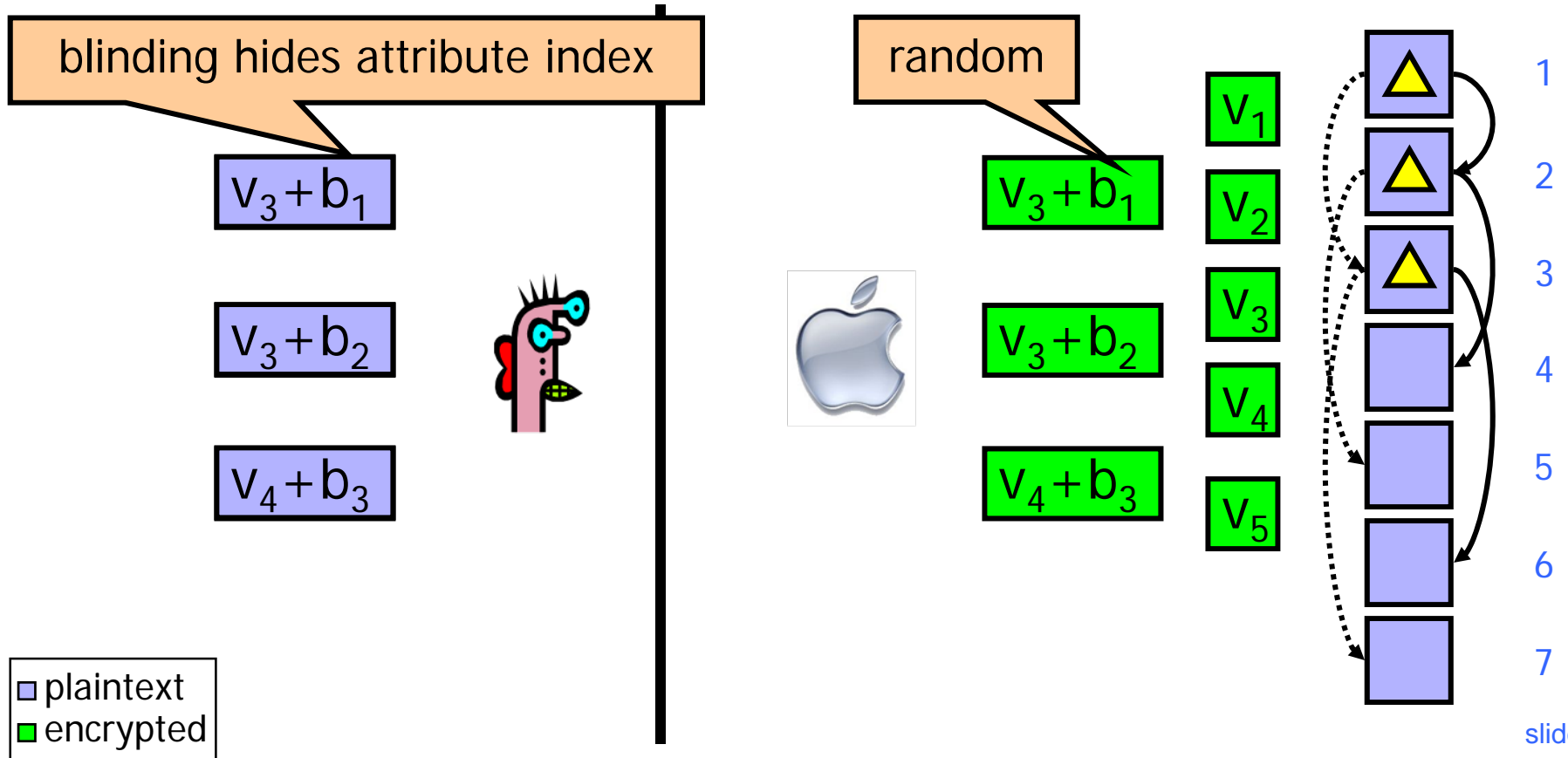


■ plaintext  
■ encrypted

# Protocol walkthrough

Hiding subset of attributes that are used

2. Vendor blinds attribute values under encryption for each decision node





# Protocol walkthrough

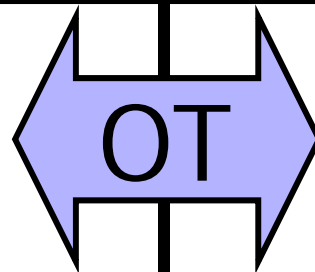
3. Blinded attribute values are converted to Yao circuit inputs using oblivious transfer

$v_3 + b_1$

$v_3 + b_2$

$v_4 + b_3$

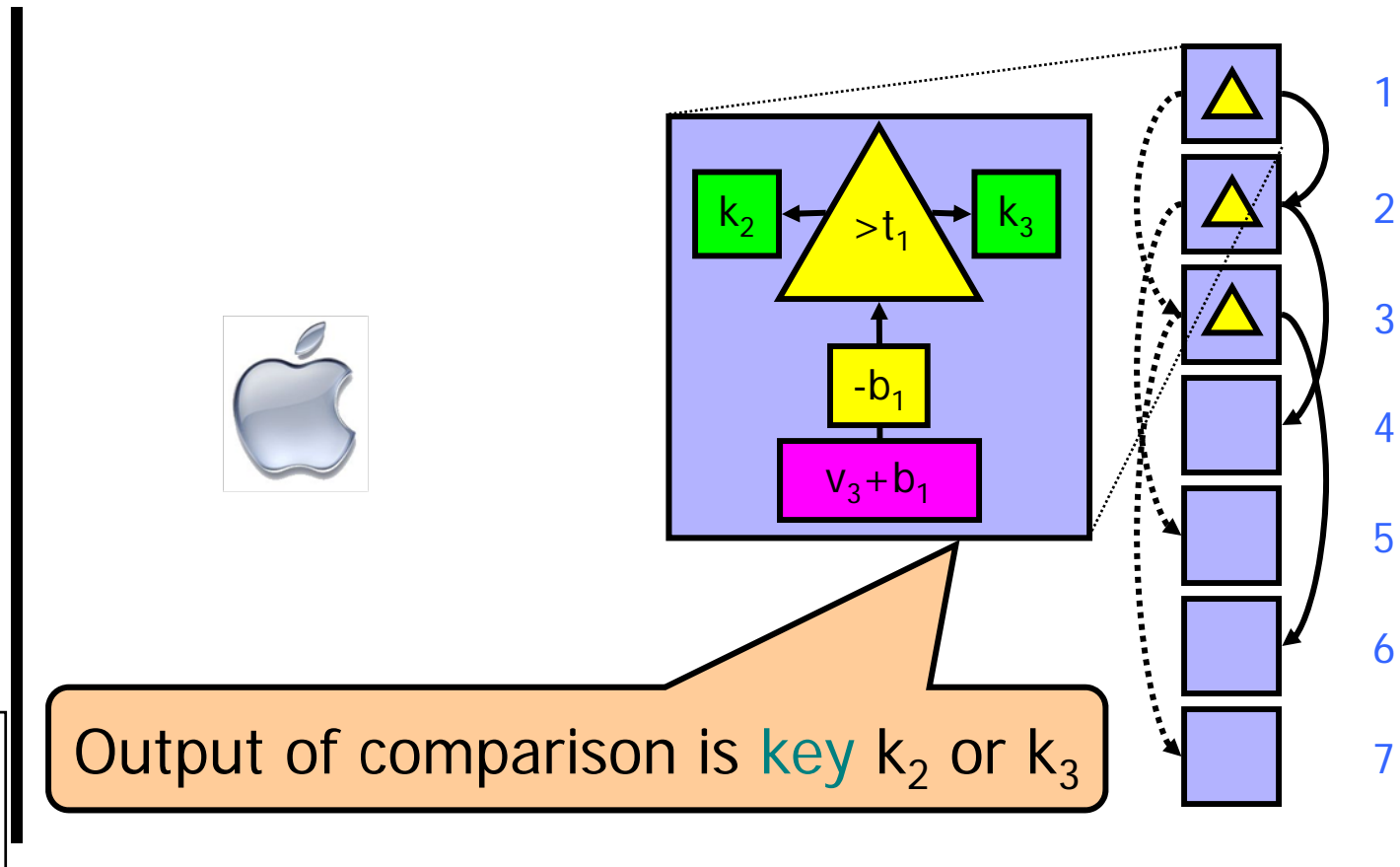
User learns Yao representation of inputs  
Vendor learns nothing



- plaintext
- encrypted
- Yao-encoded

# Protocol walkthrough

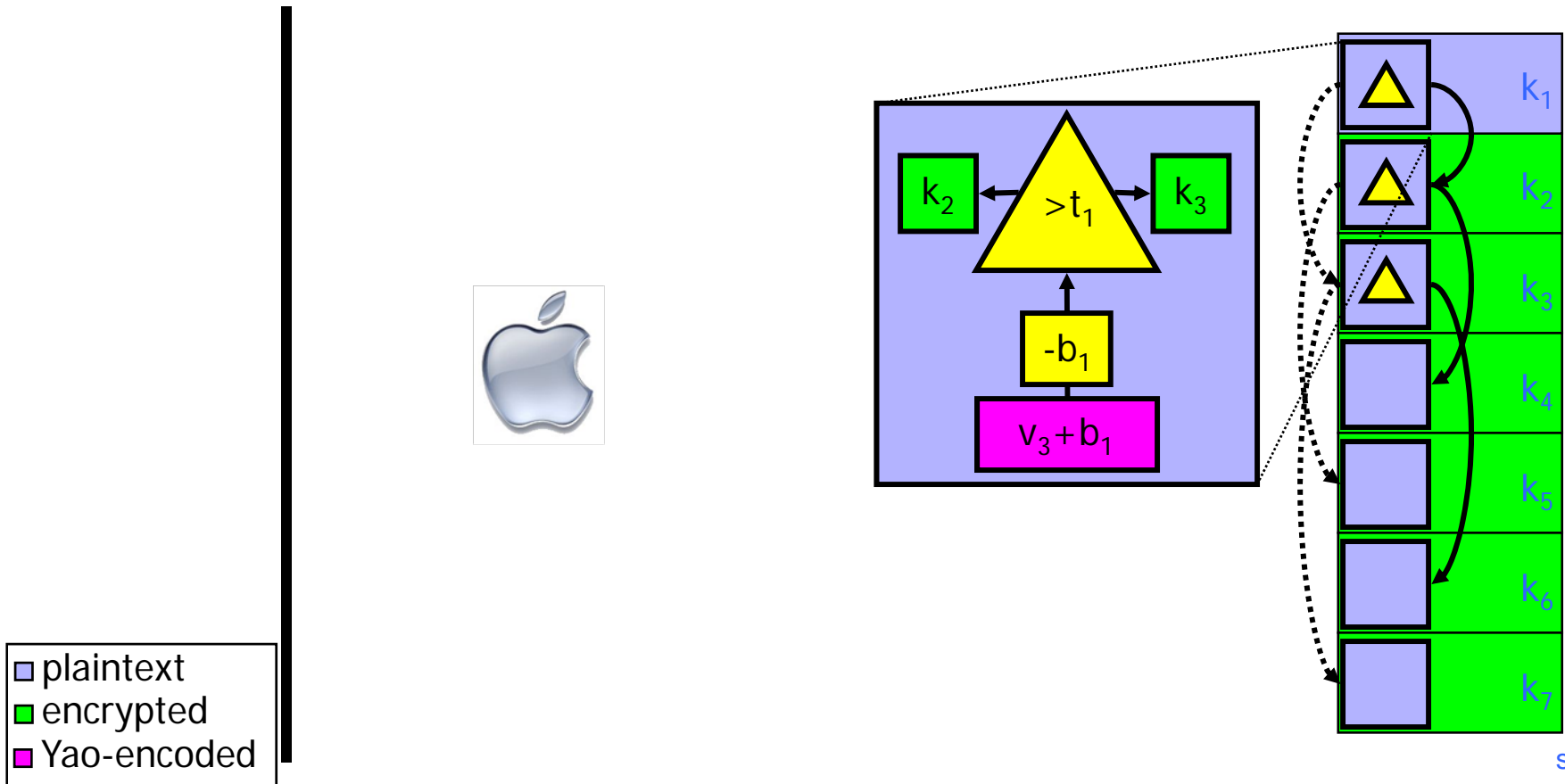
4. Vendor replaces decision nodes with secure integer comparison circuits (offline)



# Protocol walkthrough

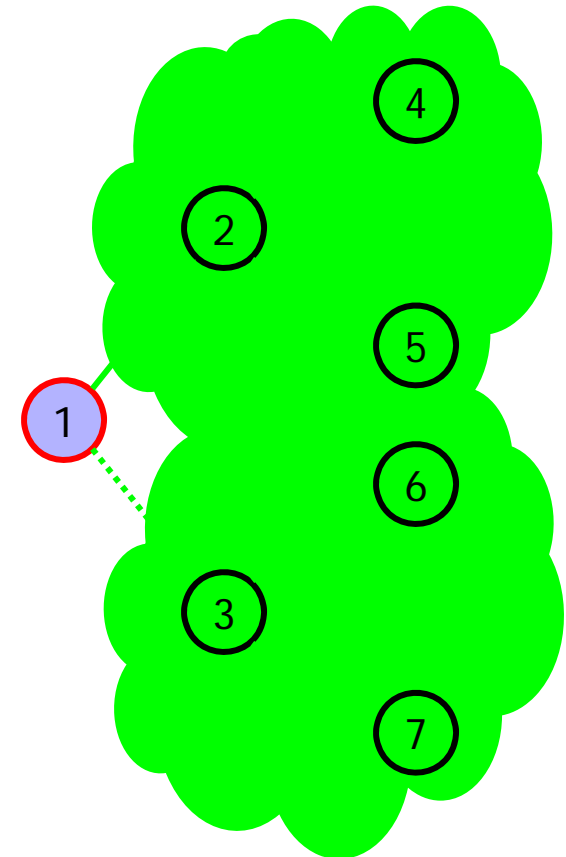
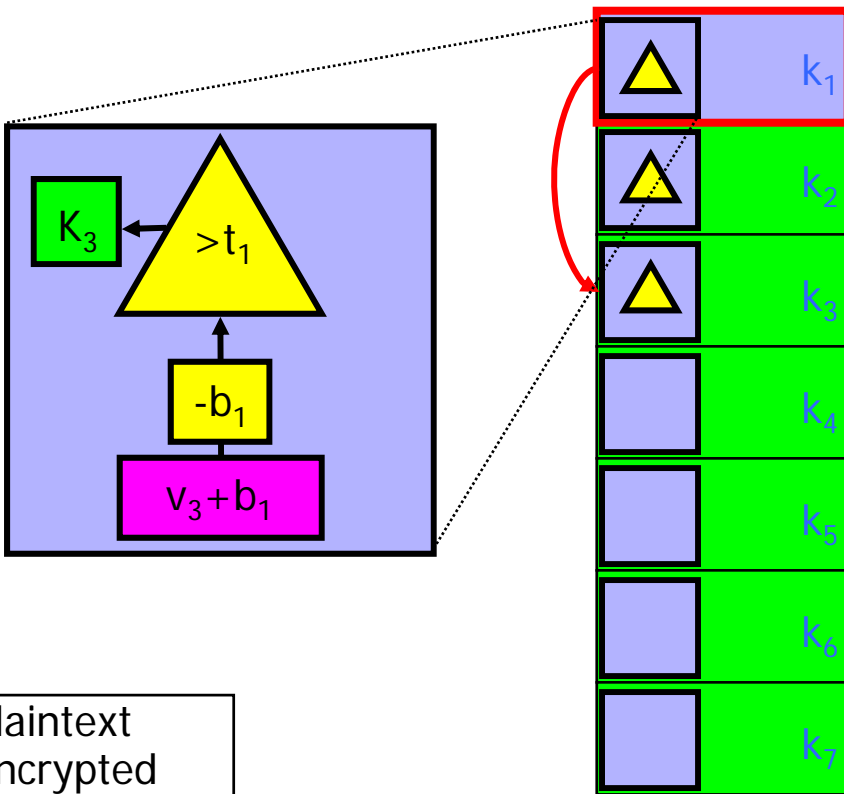
Hiding topology of T

## 5. Vendor encrypts each node



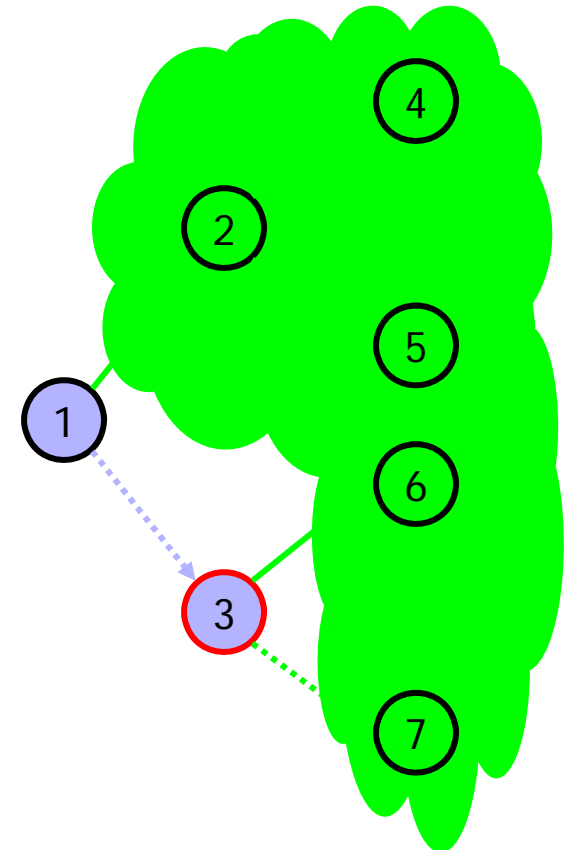
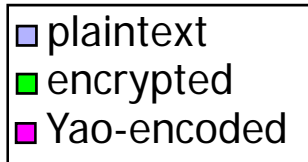
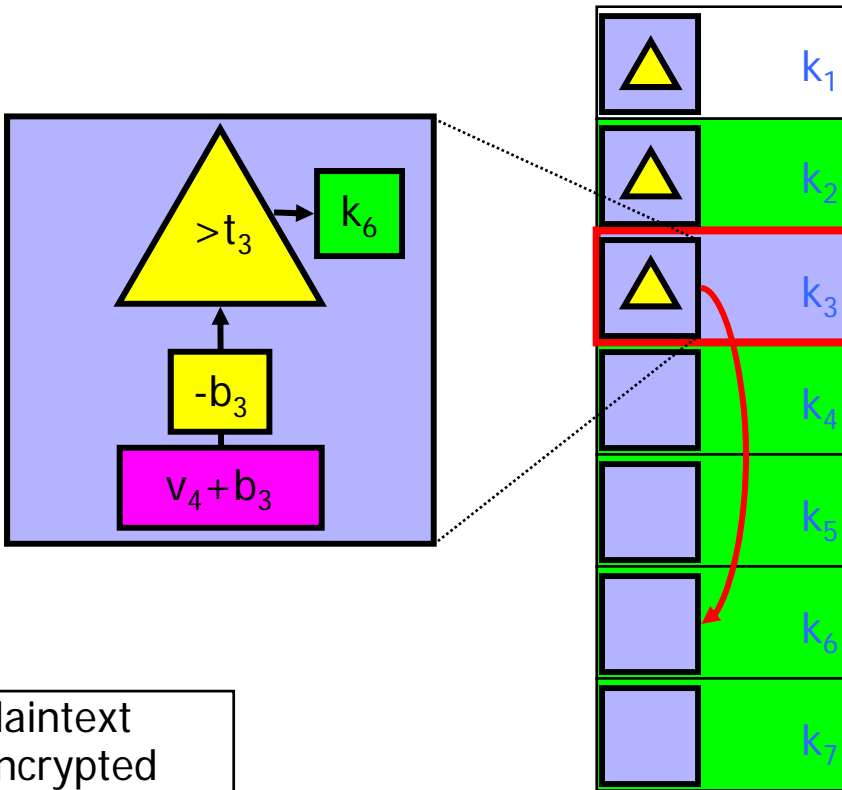
# Protocol walkthrough

## 6. User evaluates encrypted branching program



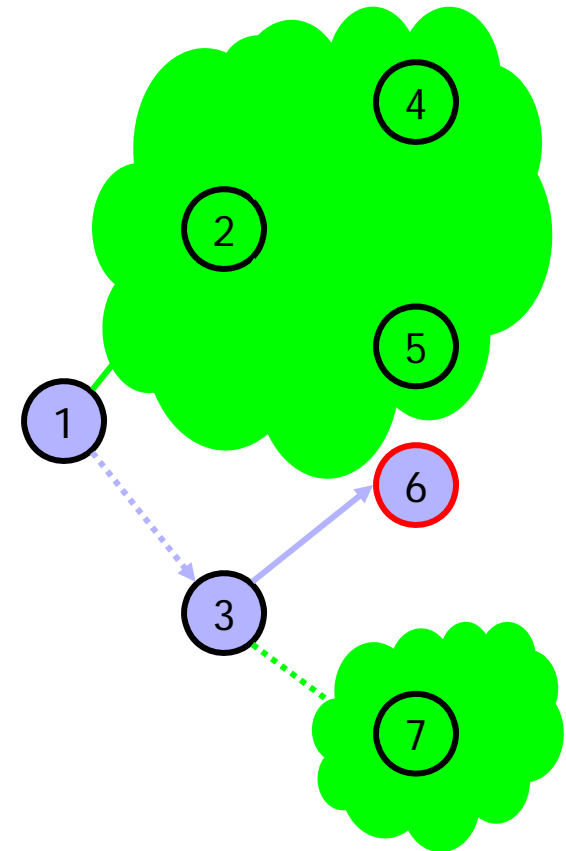
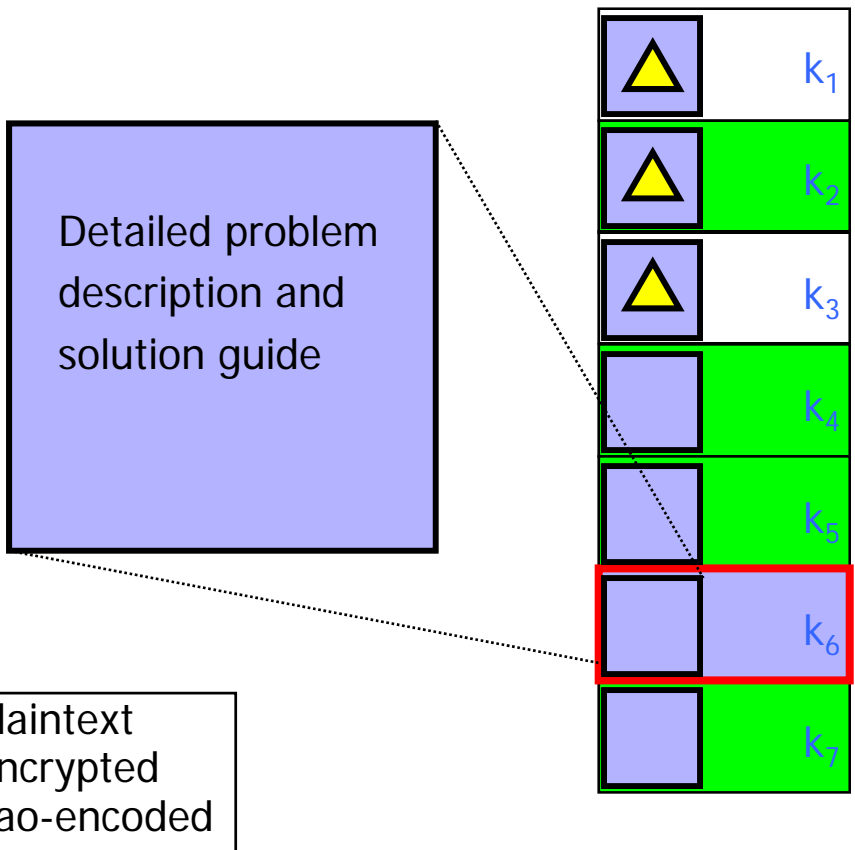
# Protocol walkthrough

## 6. User evaluates encrypted branching program



# Protocol walkthrough

## 6. User evaluates encrypted branching program



# Protocol generality

---

This is a **general** protocol

**Any** branching program, **any** attribute vector

Let's look at some specific applications

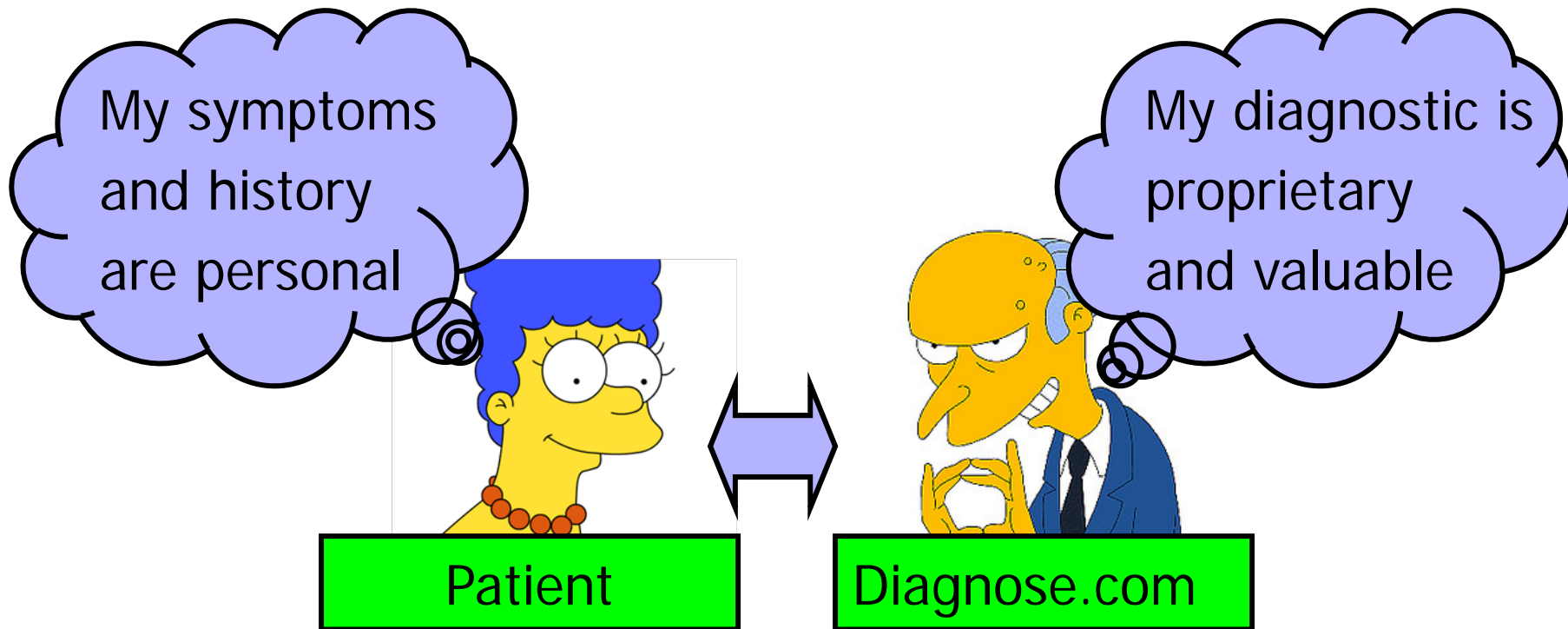
# Application: Software diagnostics

---

- ◆ Clarify [PLDI '07] automatically builds branching programs to classify anomalous program behaviors
- ◆ We have a **working system** to evaluate these branching programs securely
- ◆ The system is **practical**
  - For example, classifier for 4 cryptic gcc errors
    - 37 nodes, 2920 attributes
    - CPU: 5 seconds vendor, 7 seconds user
    - Bandwidth: 656kB vendor, 707kB user

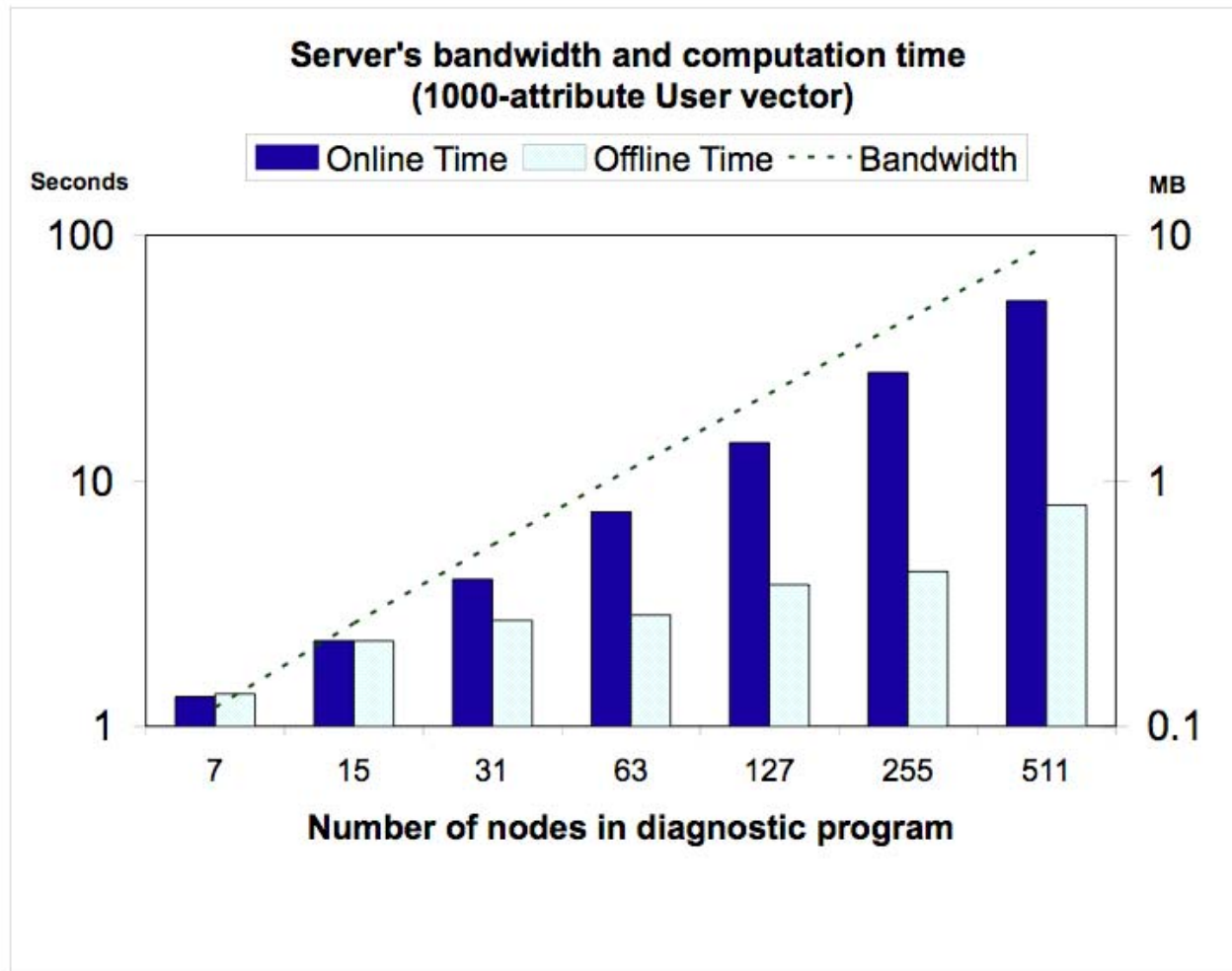


# Application: Medical diagnostics

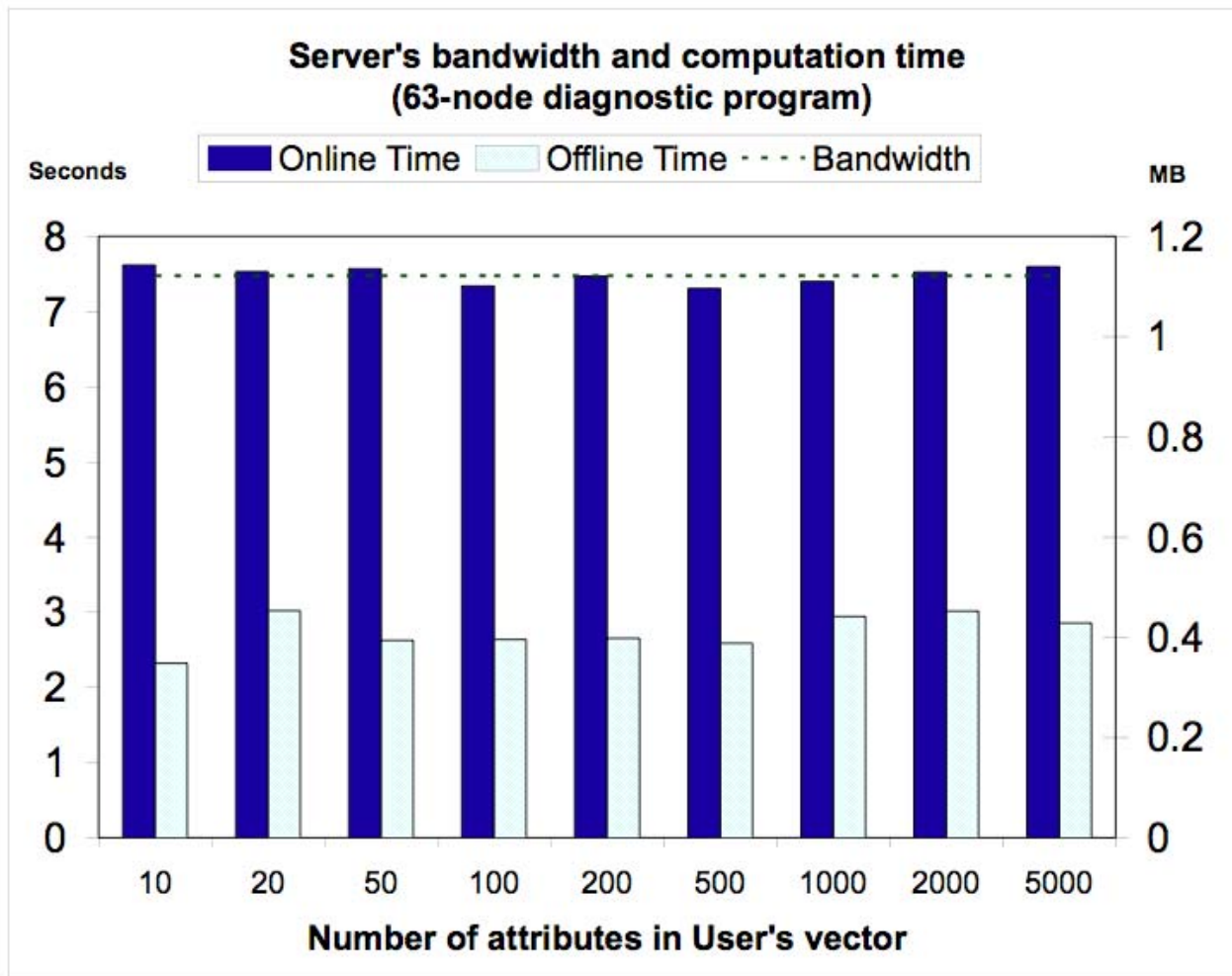


When the diagnostic  $T$  and the data  $v$  are both private, our tool can securely compute  $T(v)$

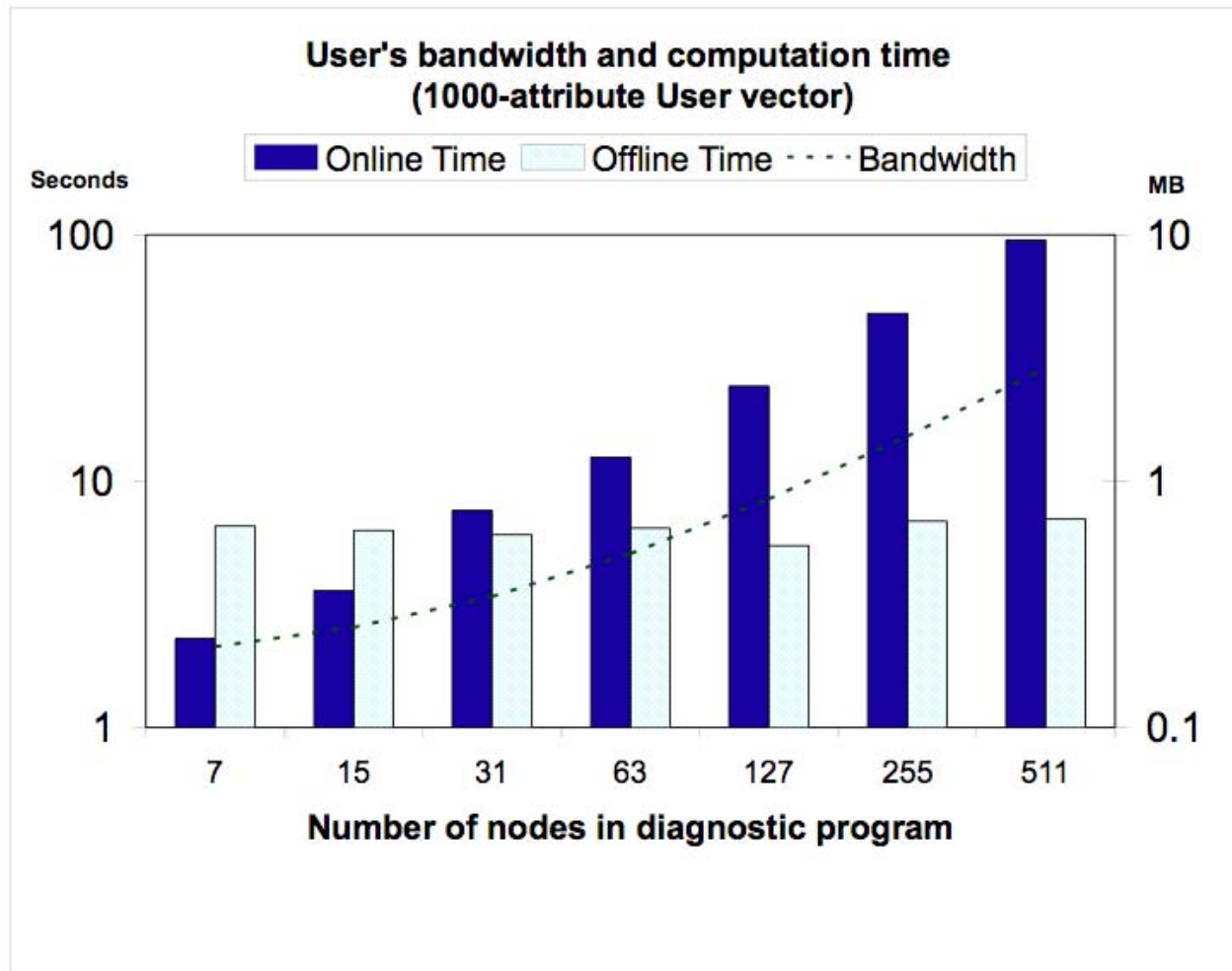
# Server performance: Size of T



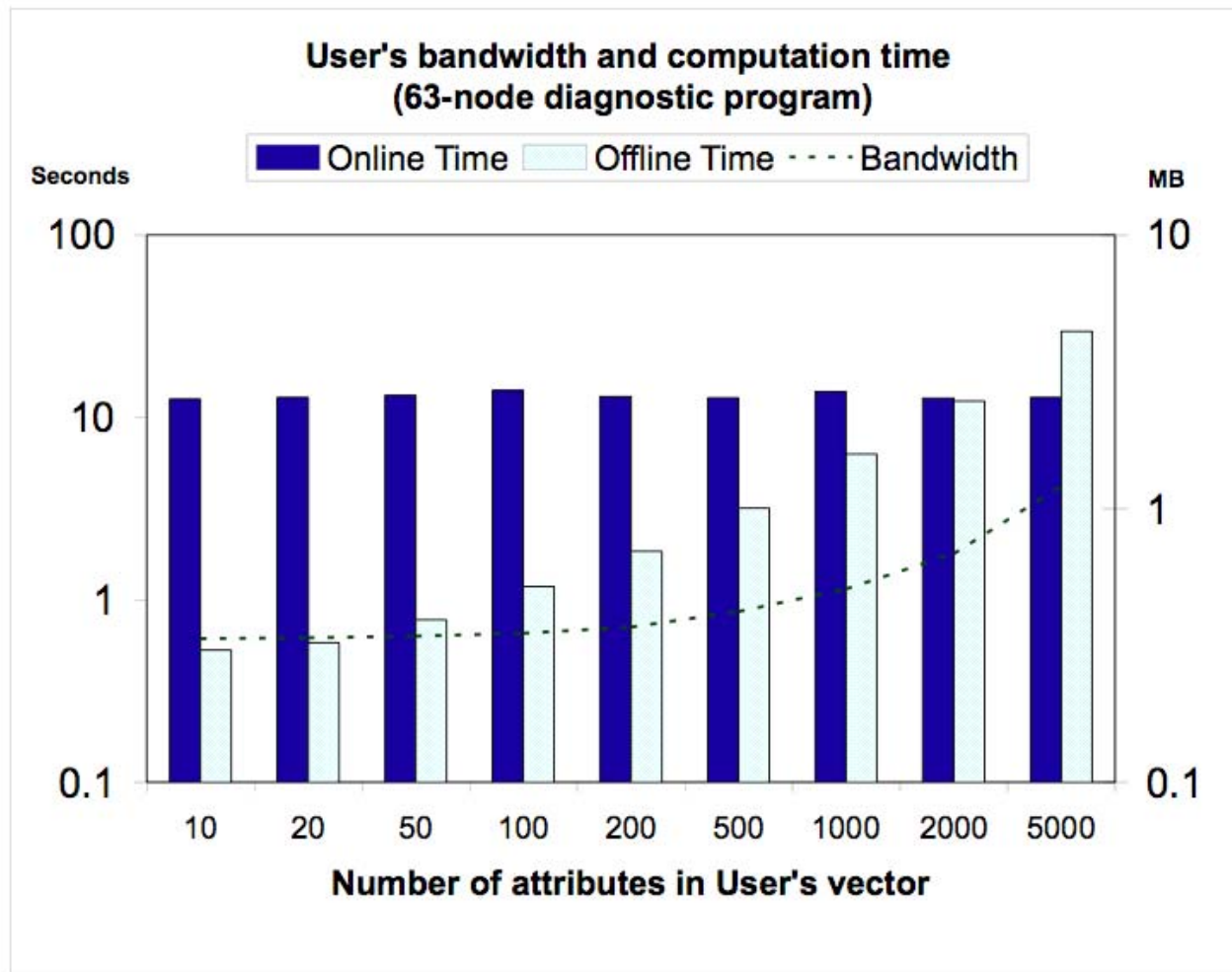
# Server performance: Size of $v$



# User performance: Size of T



# User performance: Size of $v$



# Conclusions

---

- ◆ Novel solution for secure branching program evaluation
- ◆ Provably secure
- ◆ Much more efficient than generic techniques
- ◆ Well-suited to software diagnostics
  - Online computation is independent of the size of local state
- ◆ Performance acceptable for real-world applications