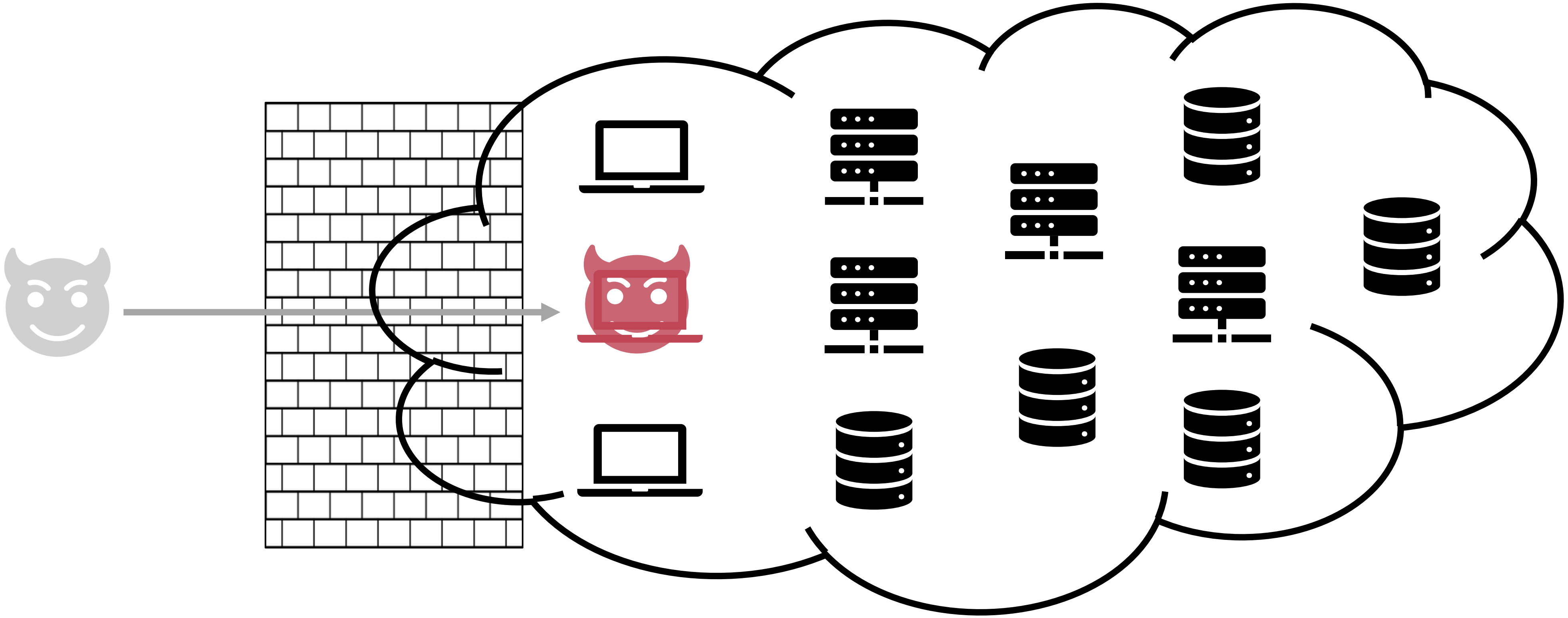
# Hopper: Modeling and Detecting Lateral Movement

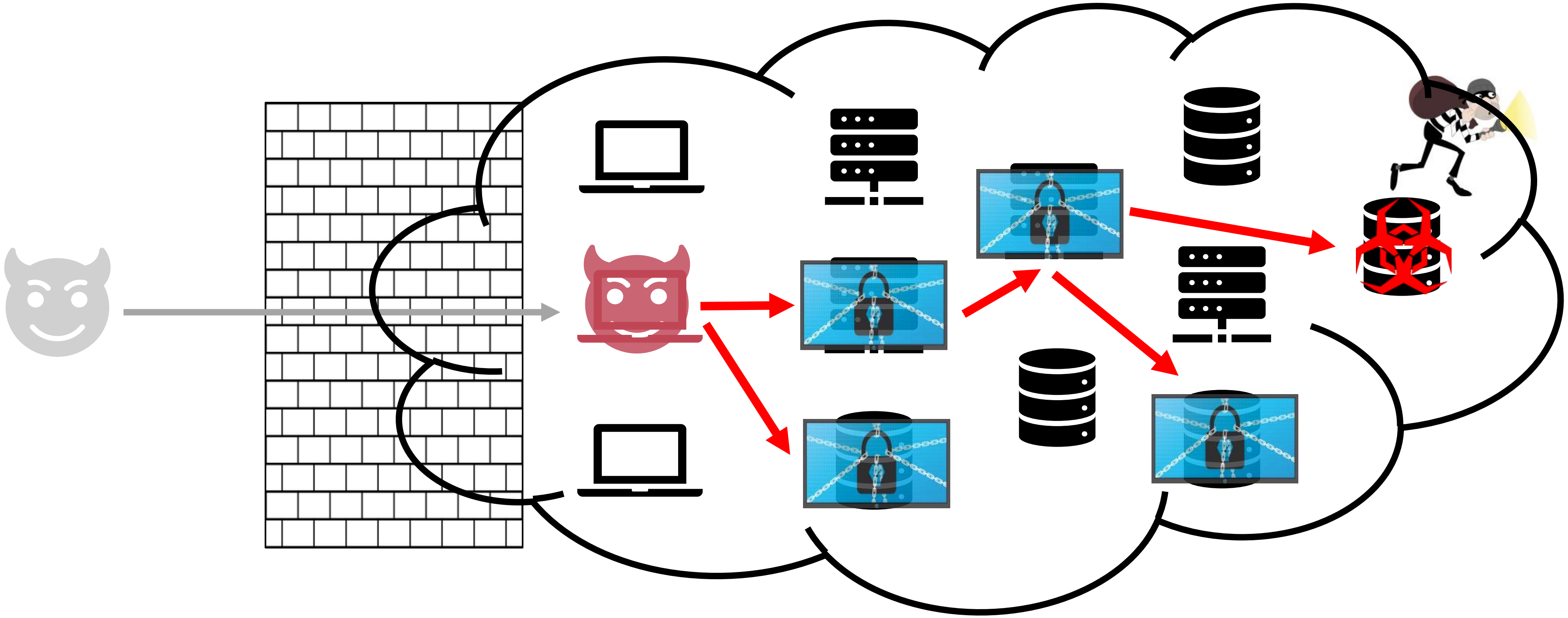
**Grant Ho***, Mayank Dhiman, Devdata Akhawe,
Vern Paxson, Geoffrey M. Voelker, Stefan Savage, David Wagner

UC San Diego*, UC Berkeley*, Dropbox*, Figma, ICSI

# How can we thwart attackers *after* they breach an enterprise's internal network?



# Enterprise attackers often need to move beyond their initial point of compromise



# Enterprise attackers often need to move beyond their initial point of compromise

USA TODAY

NEWS SPORTS ENTERTAINMENT LIFE MONEY TECH TRAVEL OPINION 50° CROSSWORDS MORE Subscri

## Timeline: North Korea and the Sony Pictures hack

USA TODAY NETWORK Lori Grisham, USA TODAY Network Published 6:39 p.m. ET Dec. 18, 2014 | Updated 12:36 p.m. ET Jan. 5, 2015



What Happened
How You May Be Affected
What You Can Do
What We Are Doing to Help

Recent Updates

## What Happened

In 2015, OPM announced **two sep**

he data of Federal gover

ıne 2015, OPM discovered

**prospective Federal emp**

agency incident response t

mation, including the Social

the background investigation databases. This includes 19.7 million individuals that applied for

ckground investigation, and 1.8 million non-applicants, primarily spouses or co-habitants of

icants. Some records also include findings from interviews conducted by background

stigators and approximately 5.6 million include fingerprints. Usernames and passwords that

ground investigation applicants used to fill out their background investigation forms were also

ın.

BBC Sign in News Sport Reel Worklife Travel Future

NEWS

Home Video World US & Canada UK Business Tech Science Stories Enter

Technology

## US hospitals turn away patients as ransomware strikes

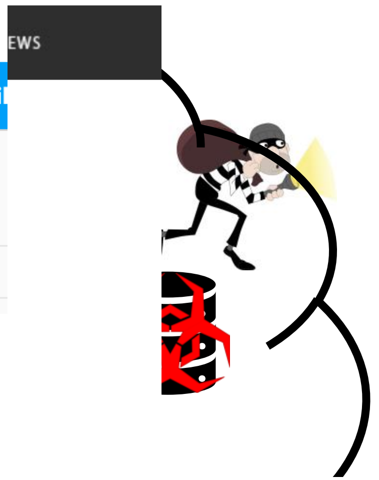
2 October 2019

Share

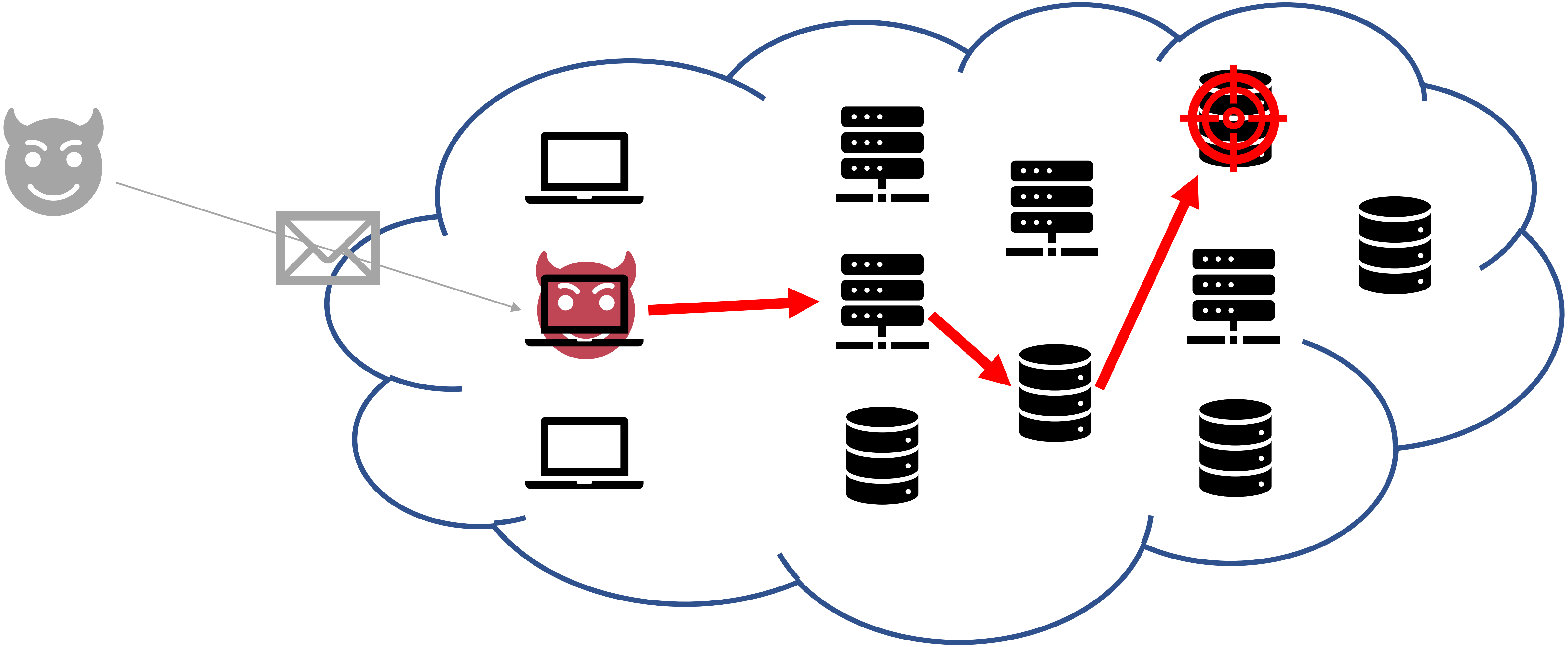
ANDY GREENBERG SECURITY 06.12.2017 08:00 AM

## 'Crash Override': The Malware That Took Down a Power Grid

**In Ukraine, researchers have found the first real-world malware that attacks physical infrastructure since Stuxnet.**



# Lateral Movement:
## Attacker movement *between* internal machines



# **The Problem:** Detecting Lateral Movement

**Threat model:** attacker has successfully compromised an internal *Machine A* and wants to get to some target *Machine Z*

**Goal:** detect malicious movement b/t internal machines w/ *low false positives*

**Prior work:** *anomalous* movement activity = an attack

- "Authentication graphs: Analyzing user behavior within an enterprise network". A Kent et al. 2015
- "Detecting Structurally Anomalous Logins Within Enterprise Networks". H Siadati, N Memon. 2017
- "Latte: Large-Scale Lateral Movement Detection". Q Liu et al. 2018
- "Log2vec: A Heterogeneous Graph Embedding Based Approach for Detecting Cyber Threats within Enterprise". Liu et al. 2019
- "Detecting Lateral Movement in Enterprise Computer Networks with Unsupervised Graph AI". B Bowman et al. 2020
- …

# **The Problem:** Detecting Lateral Movement

**Goal:** detect malicious movement between internal machines
with *low false positives*

**Prior work:** *anomalous* movement activity = an attack

**Key Limitation**: Prior state-of-the-art generates *too many FPs*
(>= ***100's*** per day)

- Deluge of anomalous-but-benign activity in modern enterprises

# Our work: Detecting Lateral Movement

**Hopper**: detects malicious movement between internal machines

- Detects > 94% attacks with < 9 FP per day
- Evaluated on 15 months of data at Dropbox
- No labeled data needed

**Key insight:** look for movement that is *suspicious*
and not just statistically anomalous

# Starting point: Internal login graph

Movement between machines (ssh, RDP, Kerberos, etc.) produces “login” records

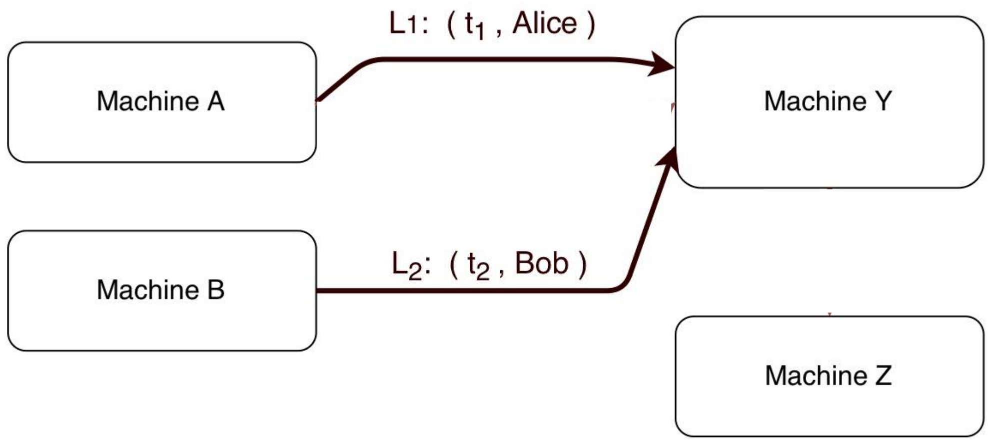

Standard login information

session start time ($t_1$), username (Alice), source machine (A), dest machine (Y)

# Detection setup: Find suspicious login paths

<u>Detection</u>

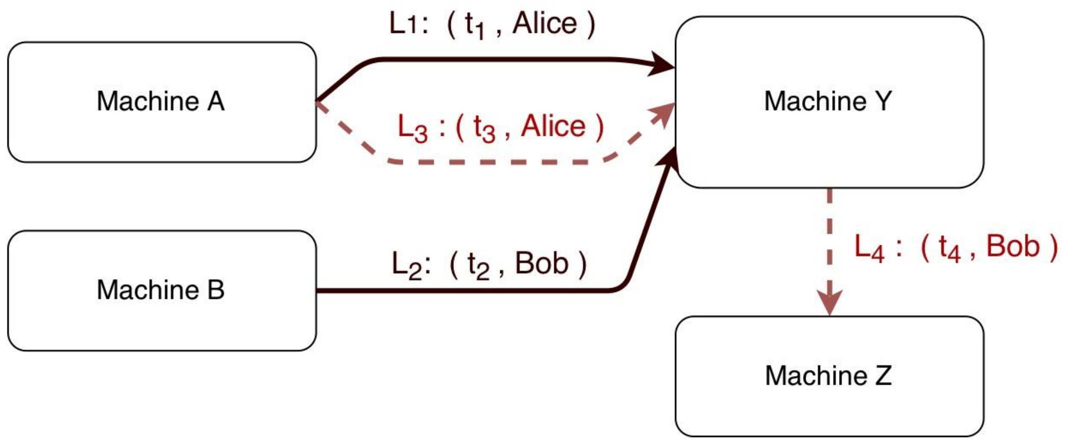
- **Training:** Build a graph from historical logins



# Detection setup: Find suspicious login paths

Detection

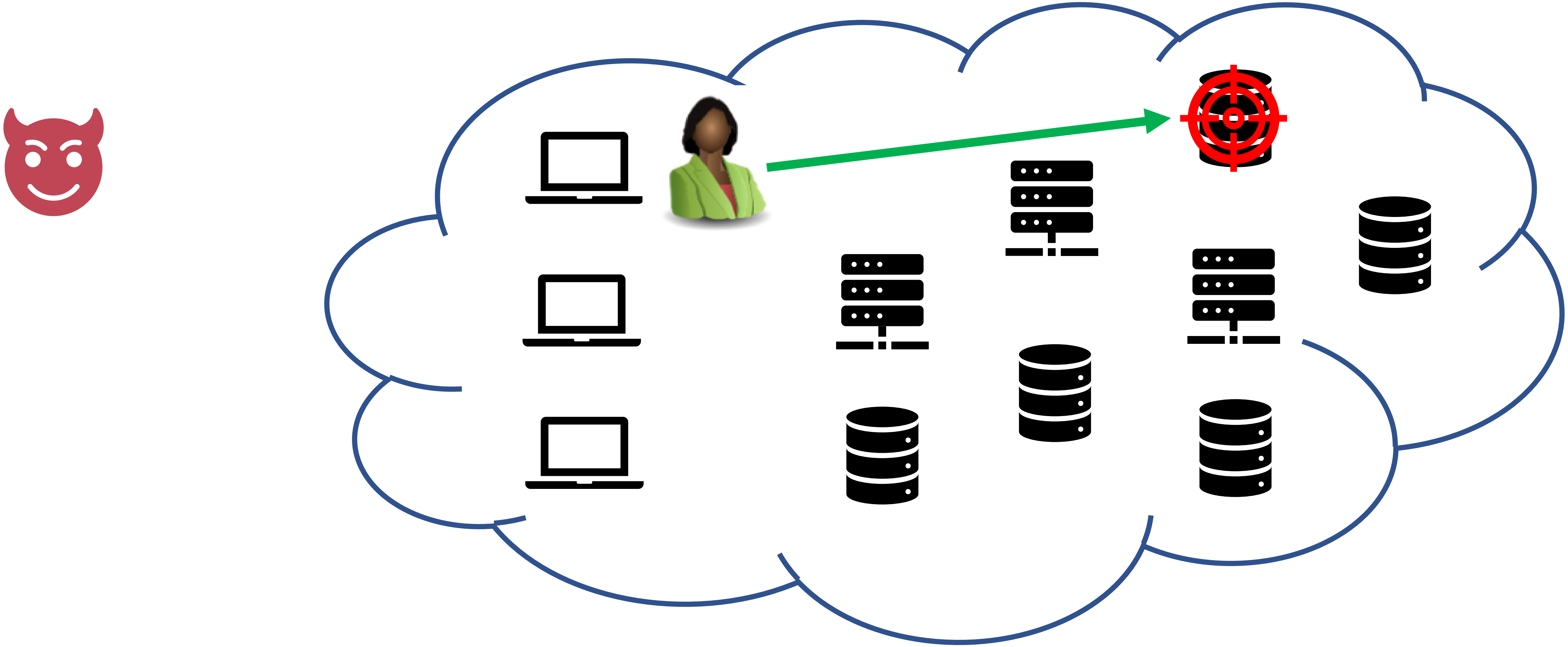
- **Training:** Build a graph from historical logins
- **Test:** Given a new set of logins, do any form a ***suspicious*** path?



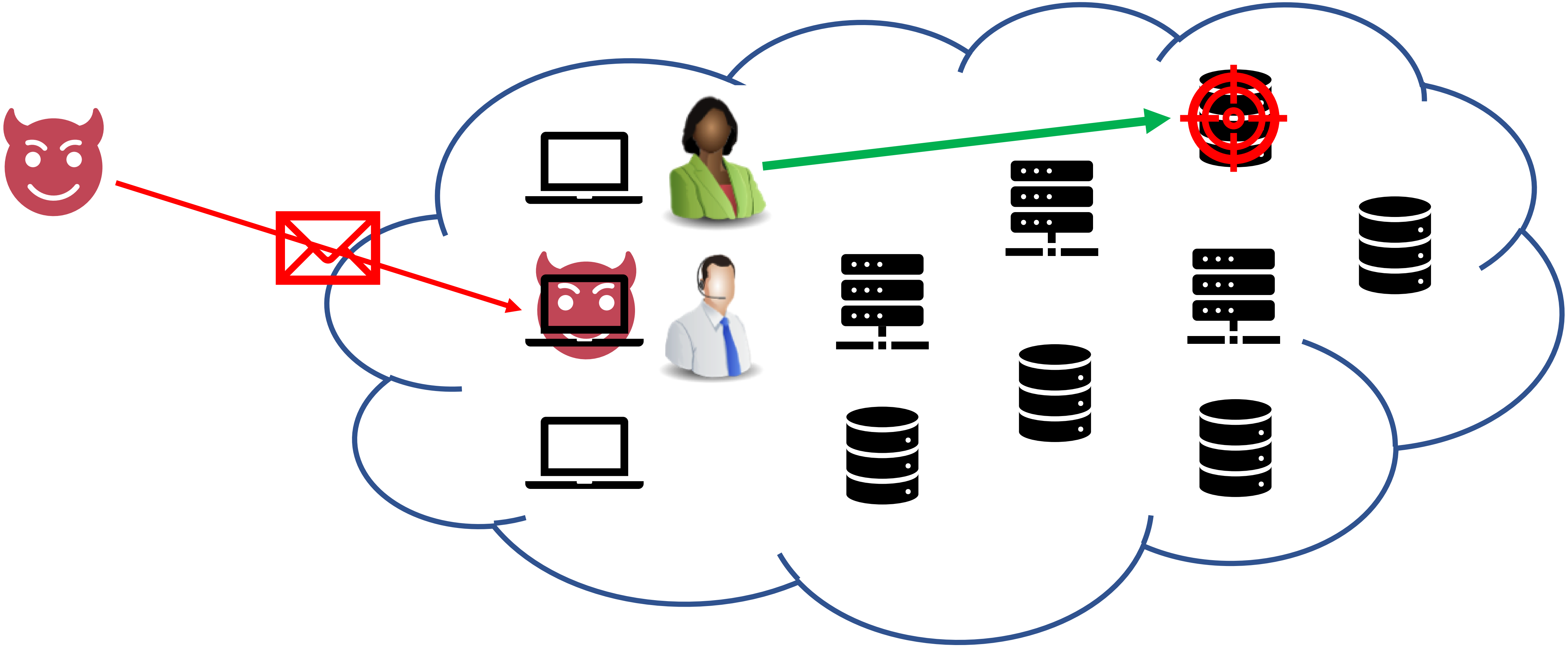

Key Question

What does it mean for a login path to be "suspicious"?

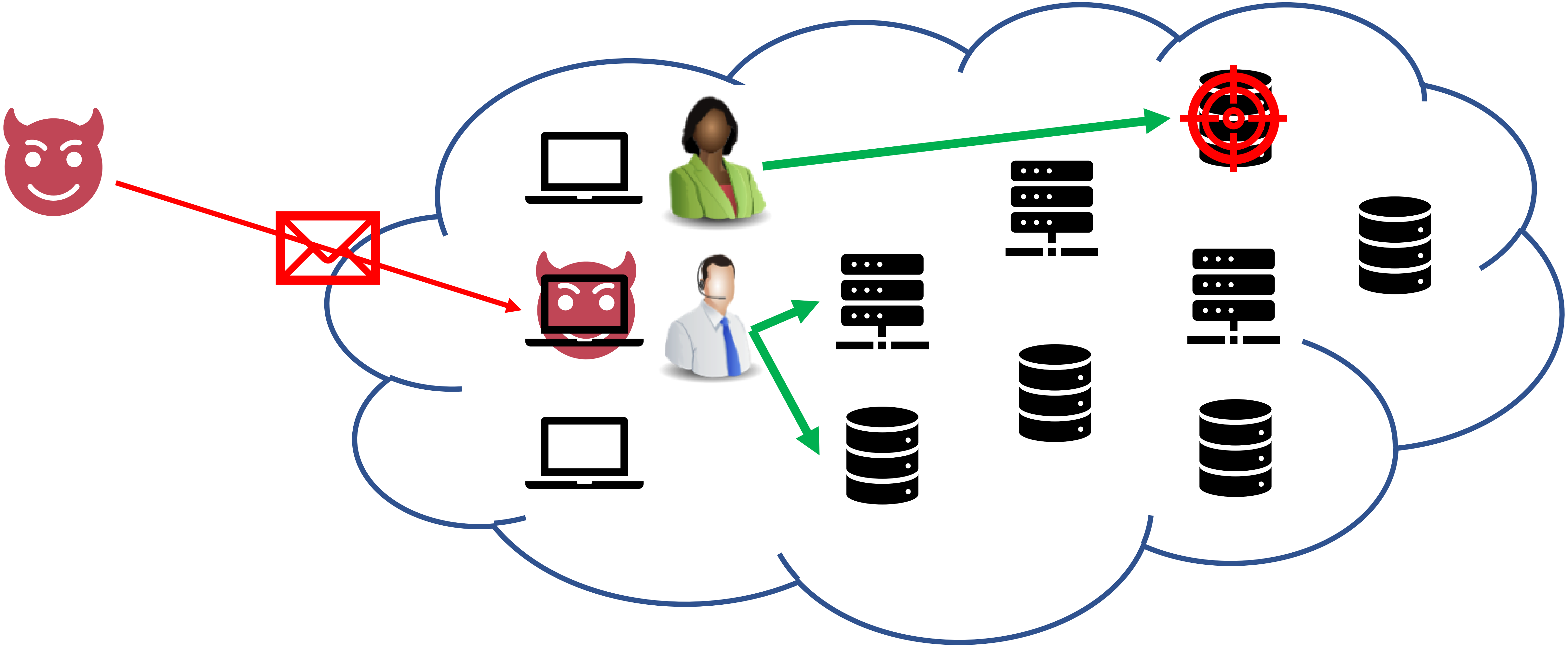
# What is a suspicious path?
# Decomposing Lateral Movement



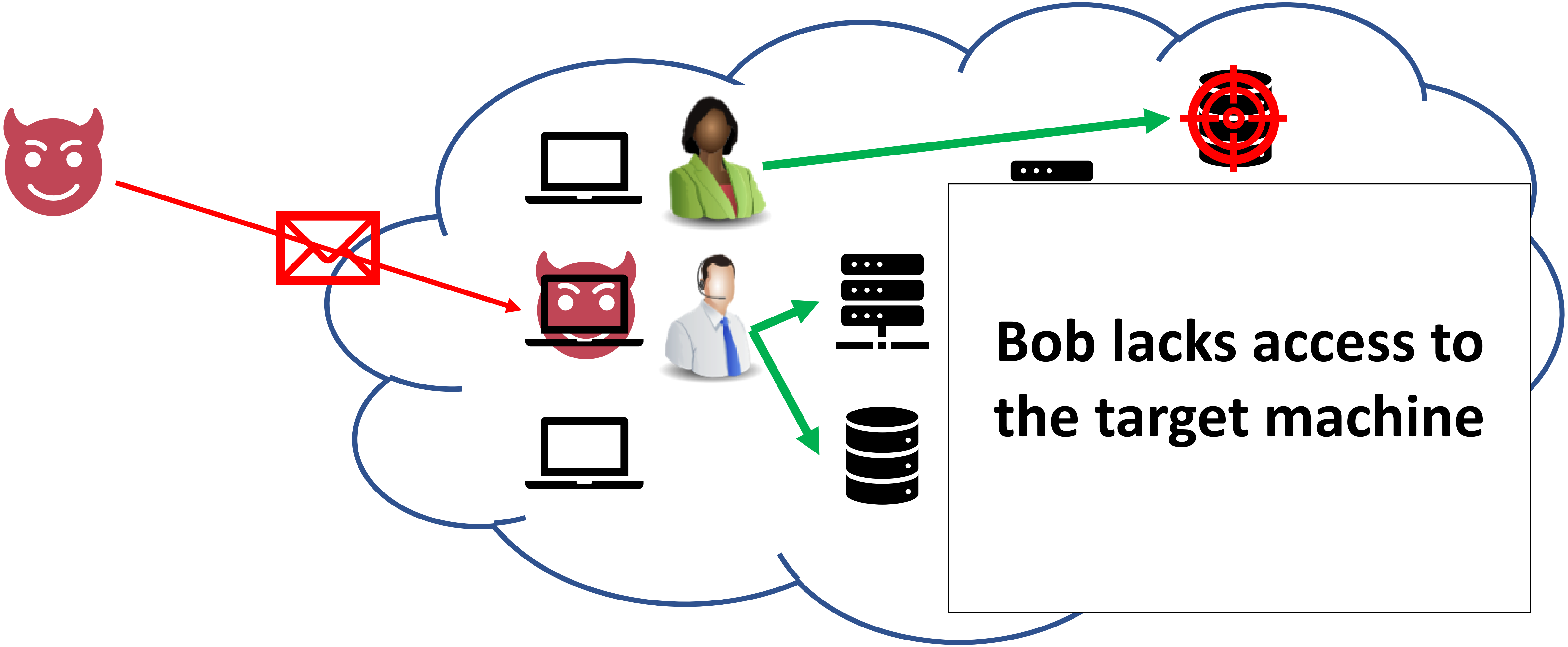
# What is a suspicious path?
# Decomposing Lateral Movement



# What is a suspicious path?
# Decomposing Lateral Movement

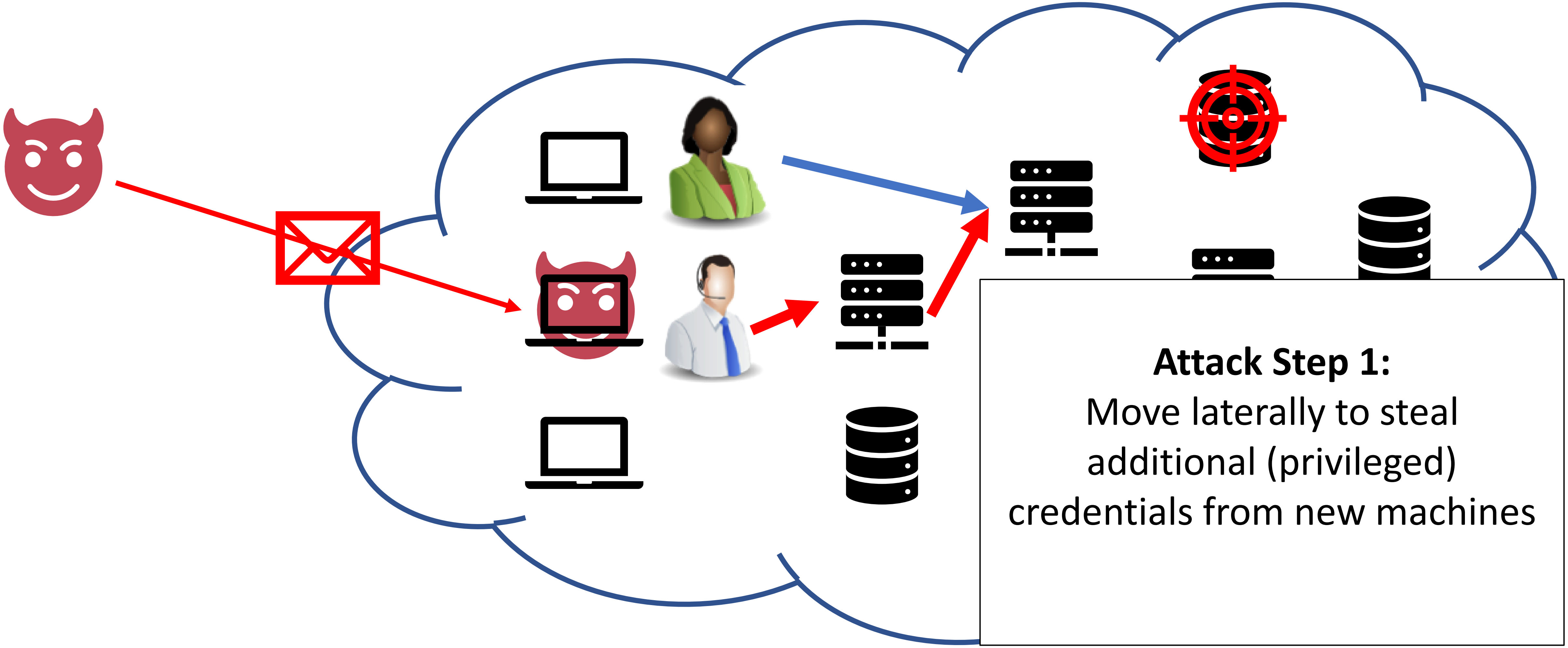


# What is a suspicious path?
# Decomposing Lateral Movement



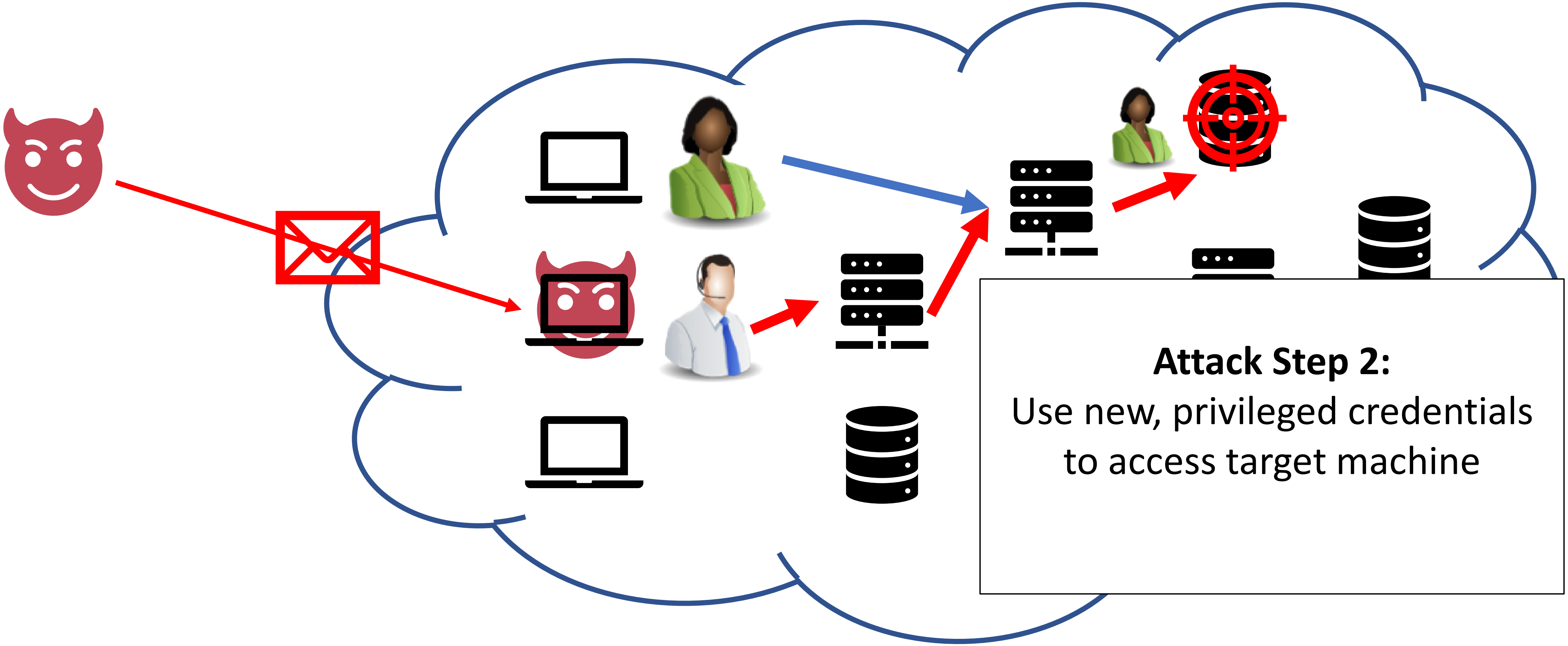
**Bob lacks access to the target machine**

# What is a suspicious path?
# Decomposing Lateral Movement



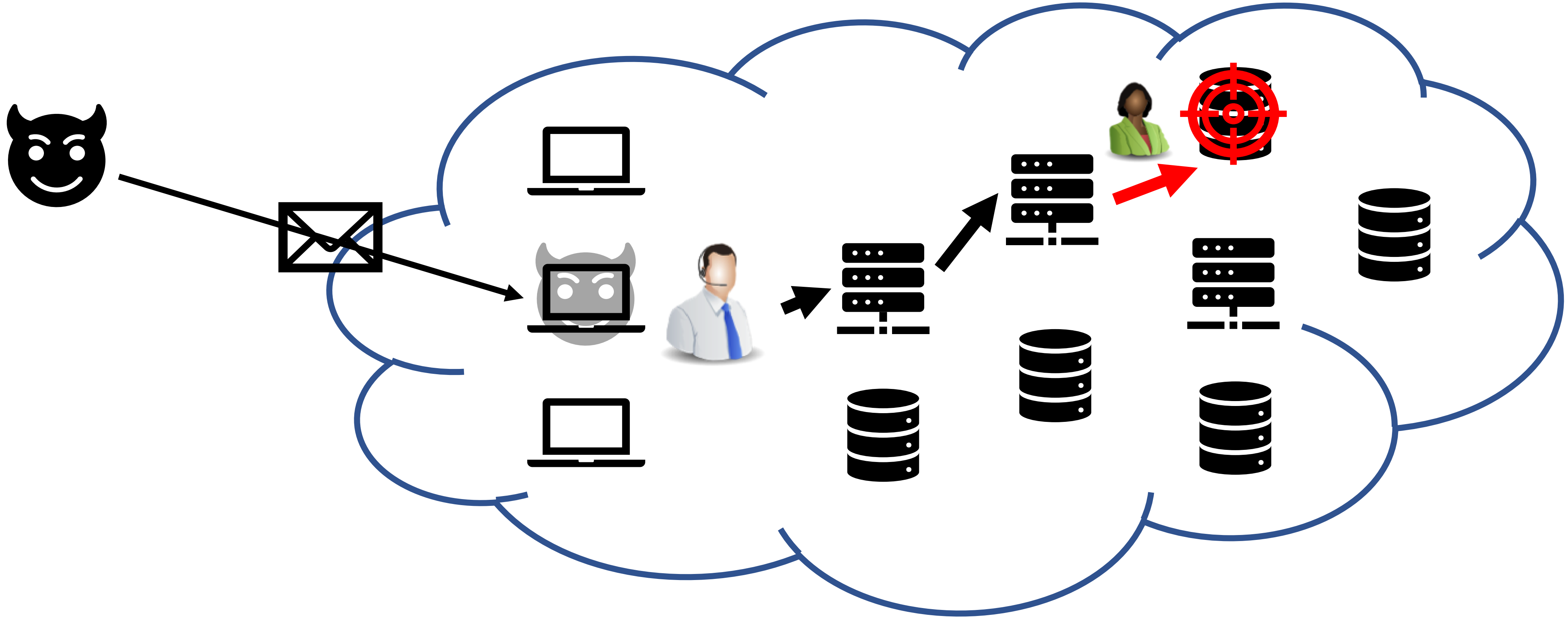
**Attack Step 1:**
Move laterally to steal additional (privileged) credentials from new machines

# What is a suspicious path?
# Decomposing Lateral Movement



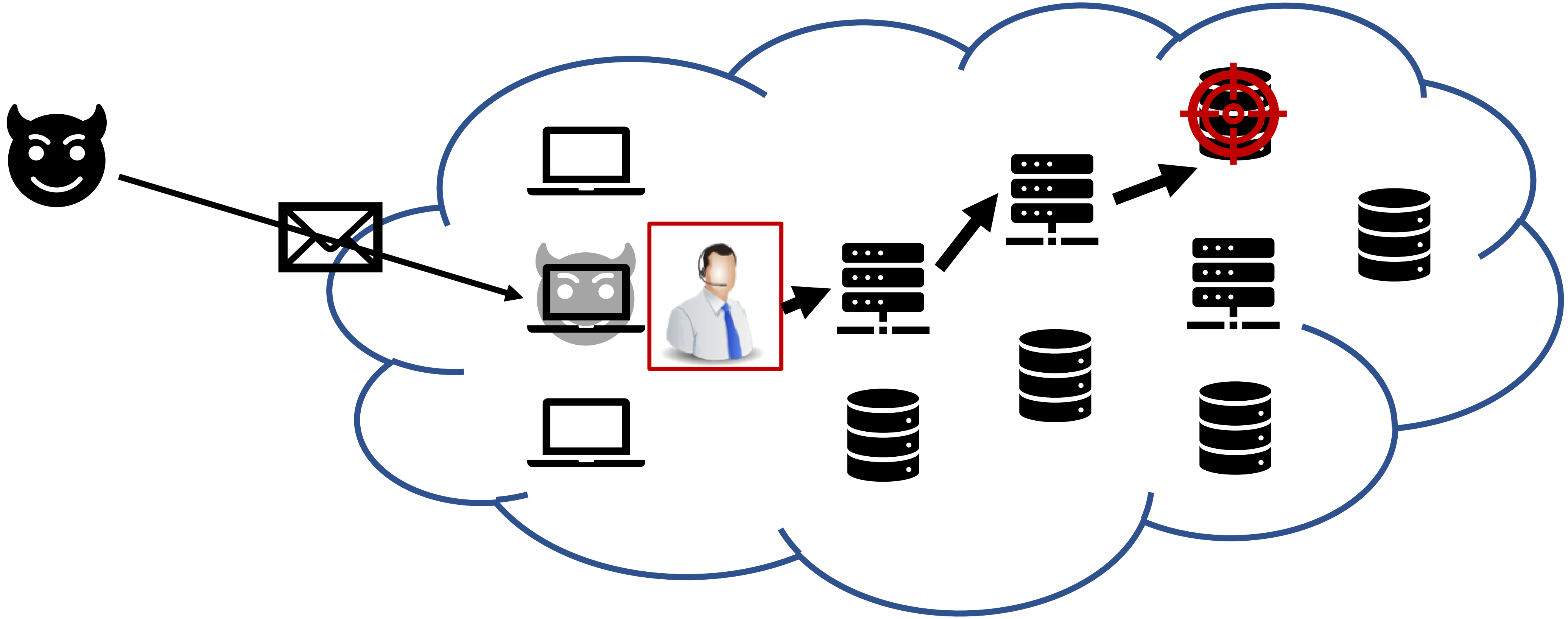
**Attack Step 2:**
Use new, privileged credentials to access target machine

# Lateral Movement paths: 2 suspicious properties



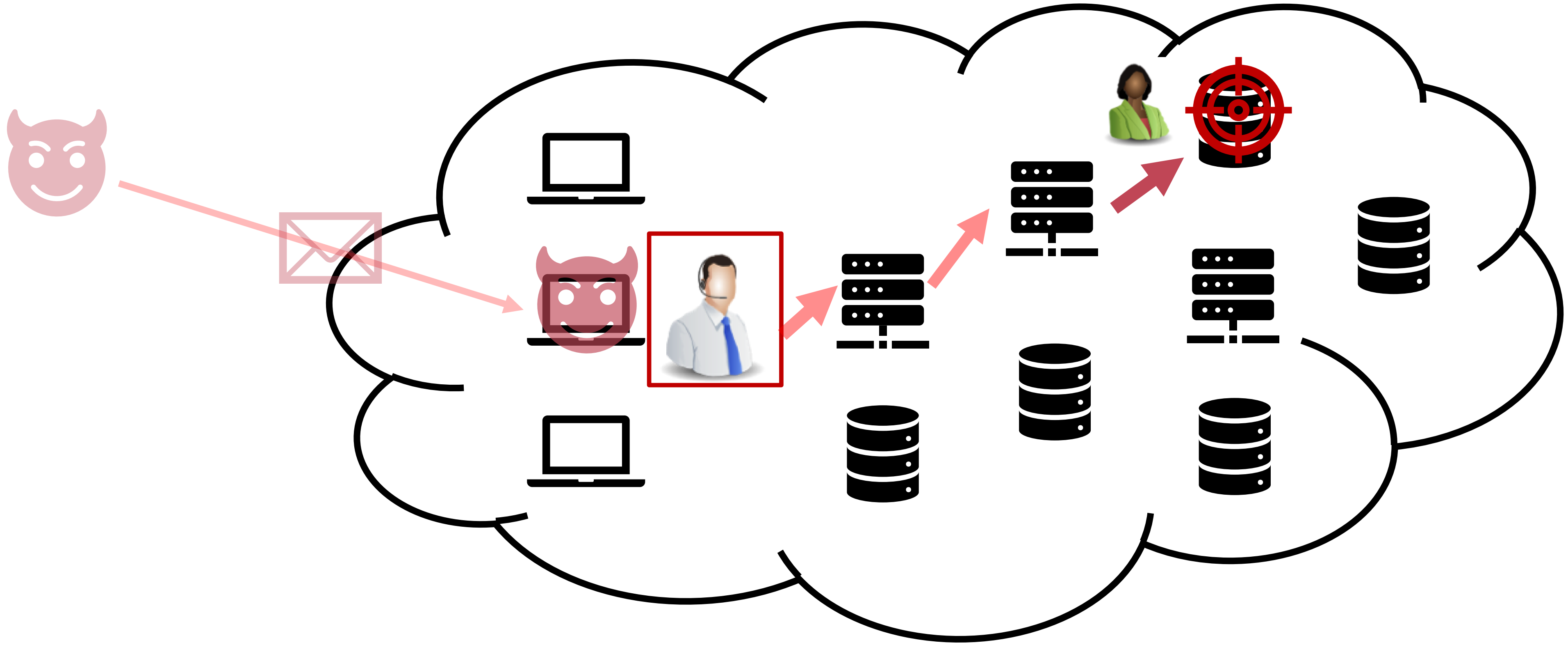
Property #1: path contains 1+ login that
uses a new or unexpected set of credentials

# Lateral Movement paths: 2 suspicious properties



Property #2: path accesses a machine that the initial user does not have legitimate access to

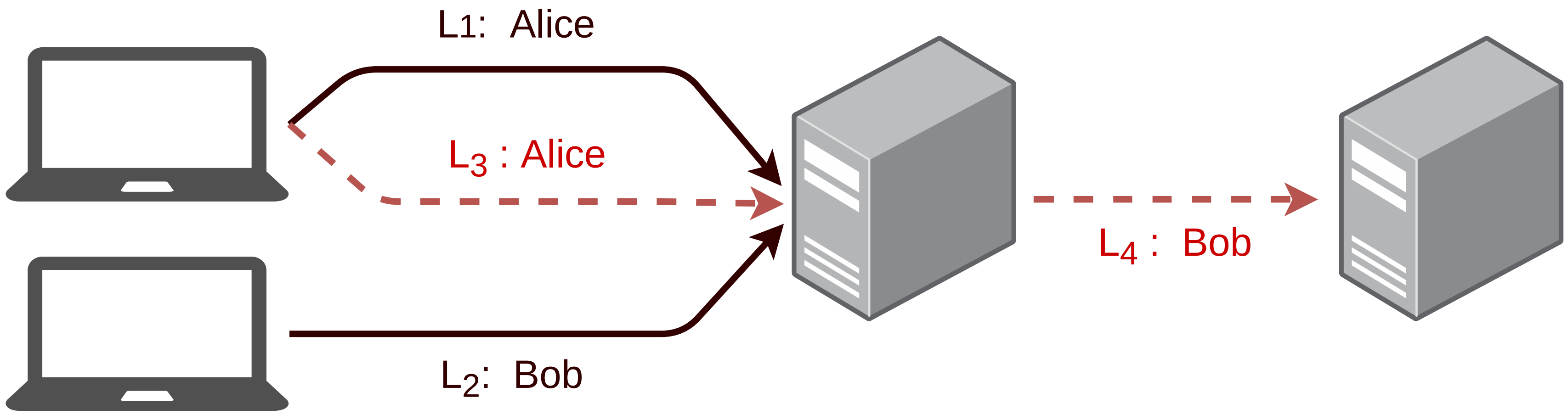
# Hopper: Identifying suspicious paths: 2 key properties



**Property #1:** path has a login that uses an unexpected set of credentials
**Property #2:** path accesses a machine that the initial user could not access

# Overview: Key sub-problems + our solutions

**Correctly identifying which set of logins form paths "caused" by same user**




- Which inbound login forms a path with login $L_4$?
  - Real-world authentication logs don't provide causality information

# Overview: Key sub-problems + our solutions

**Correctly identifying which set of logins form paths “caused” by same user**

- Methods to infer login causality using enterprise domain knowledge

**Handling gaps & ambiguity in path inference**

- Conservatively infer multiple potential paths
- Specification-based anomaly detection:

  reduce FP by *selectively* applying anomaly detection

  only to paths that potentially contain both suspicious properties

# Evaluation

15 months of data from Dropbox's internal corp network: 700M+ logins

- 1 red-team attack + 326 simulated attacks :
  various goals (e.g., ransomware & targeted compromise) + stealthiness

| | Hopper |
|---|---|
| **True Positives (Detection Rate)** | **309 / 327** |
| **False Positives** | **3,560** |
| **Avg Daily Alerts** | **9 alerts / day** |

# Evaluation

15 months of data from Dropbox's internal corp network: 700M+ logins

- 1 red-team attack + 326 simulated attacks :
  various goals (e.g., ransomware & targeted compromise) + stealthiness

| | Hopper | SAL (CCS 2017) Equal Detection |
|---|---|---|
| **True Positives (Detection Rate)** | **309 / 327** | 309 / 327 |
| **False Positives** | **3,560** | 27,927 |
| **Avg Daily Alerts** | **9 alerts / day** | 71 alerts / day |

# Evaluation

15 months of data from Dropbox's internal corp network: 700M+ logins

- 1 red-team attack + 326 simulated attacks :
  various goals (e.g., ransomware & targeted compromise) + stealthiness

| | Hopper | SAL (CCS 2017) Equal Detection |
|---|---|---|
| **True Positives (Detection Rate)** | **309 / 327** | 309 / 327 |
| **False Positives** | **3,560** | 27,927 |
| **Avg Daily Alerts** | **9 alerts / day** | 71 alerts / day |

Our Work (Hopper)

- **8x** improvement over state-of-the-art (traditional anomaly detection)
- Key improvement = look for paths with suspicious structure, rather than just statistical anomalies

# Summary

- Analyzing network movement between *internal machines* can help mitigate enterprise attacks
- Enterprises have lots of anomalous-but-benign activity: need to combine anomaly detection w/ *suspicious structure* for practical detection
- Identifying *causally-related movement* is challenging, but provides a powerful detection paradigm
- Hopper, an approach built on these ideas, detected > 94% of lateral movement scenarios with < 9 FP / day across 15 months at Dropbox

**Thank you!**

**grantho@eng.ucsd.edu**