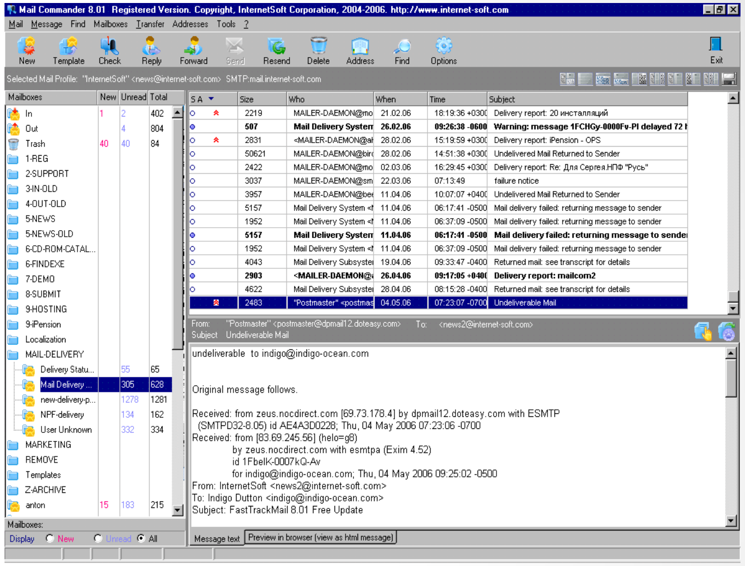
Making Programs Forget: Enforcing Lifetime for Sensitive Data

Jayanthkumar Kannan (Google Inc), **Gautam Altekar** (UC Berkeley), Petros Maniatis (Intel Labs), Byung-Gon Chun (Intel Labs)

The Problem: Lingering Data

Sensitive Data



Mail Commander 8.01 Registered Version. Copyright, InternetSoft Corporation, 2004-2006. http://www.internet-soft.com
Mail Message Find Mailboxes Transfer Addresses Tools ?
New
Template
Check
Reply
Forward
Send
Resend
Delete
Address
Find
Options
Exit

Selected Mail Profile: "InternetSoft" <news@internet-soft.com> SMTP:mail.internet-soft.com

Mailboxes	New	Unread	Total
In	1	2	402
Out		4	804
Trash	40	40	84
1-REG			
2-SUPPORT			
3-IN-OLD			
4-OUT-OLD			
5-NEWS			
5-NEWS-OLD			
6-CD-ROM-CATAL...			
6-FINDEXE			
7-DEMO			
8-SUBMIT			
9-HOSTING			
9-iPension			
Localization			
MAIL-DELIVERY			
Delivery Statu...		55	65
Mail Delivery ...		305	628
new-delivery-p...		1278	1281
NPF-delivery		134	162
User Unknown		332	334
MARKETING			
REMOVE			
Templates			
Z-ARCHIVE			
anton	15	183	215

Mailboxes:
Display ○ New ○ Unread ● All

S A	Size	Who	When	Time	Subject
○ ⇑	2219	MAILER-DAEMON@mo	21.02.06	18:19:36 +0300	Delivery report: 20 инсталляций
●	**507**	**Mail Delivery System**	**26.02.06**	**09:26:38 -0600**	**Warning: message 1FCHGy-0000Fv-PI delayed 72 h**
○ ⇑	2831	<MAILER-DAEMON@al	28.02.06	15:19:59 +0300	Delivery report: iPension - OPS
○	50621	MAILER-DAEMON@bird	28.02.06	14:51:38 +0300	Undelivered Mail Returned to Sender
○	2422	MAILER-DAEMON@mo	02.03.06	16:29:45 +0300	Delivery report: Re: Для Сергея.НПФ "Русь"
○	3037	MAILER-DAEMON@sm	22.03.06	07:13:49	failure notice
○	3957	MAILER-DAEMON@bee	11.04.06	10:07:07 +0400	Undelivered Mail Returned to Sender
○	5157	Mail Delivery System <I	11.04.06	06:17:41 -0500	Mail delivery failed: returning message to sender
○	1952	Mail Delivery System <I	11.04.06	06:37:09 -0500	Mail delivery failed: returning message to sender
●	**5157**	**Mail Delivery System**	**11.04.06**	**06:17:41 -0500**	**Mail delivery failed: returning message to sender**
○	1952	Mail Delivery System <I	11.04.06	06:37:09 -0500	Mail delivery failed: returning message to sender
○	4043	Mail Delivery Subsyster	19.04.06	09:33:47 -0400	Returned mail: see transcript for details
●	**2903**	**<MAILER-DAEMON@a**	**26.04.06**	**09:17:05 +0400**	**Delivery report: mailcom2**
○	4622	Mail Delivery Subsyster	28.04.06	08:15:28 -0400	Returned mail: see transcript for details
⇑	2483	"Postmaster" <postmas	04.05.06	07:23:07 -0700	Undeliverable Mail

From: "Postmaster" <postmaster@dpmail12.doteasy.com> To: <news2@internet-soft.com>
Subject Undeliverable Mail

```
undeliverable  to indigo@indigo-ocean.com


Original message follows.

Received: from zeus.nocdirect.com [69.73.178.4] by dpmail12.doteasy.com with ESMTP
  (SMTPD32-8.05) id AE4A3D0228; Thu, 04 May 2006 07:23:06 -0700
Received: from [83.69.245.56] (helo=g8)
          by zeus.nocdirect.com with esmtpa (Exim 4.52)
          id 1FbelK-0007kQ-Av
          for indigo@indigo-ocean.com; Thu, 04 May 2006 09:25:02 -0500
From: InternetSoft <news2@internet-soft.com>
To: Indigo Dutton <indigo@indigo-ocean.com>
Subject: FastTrackMail 8.01 Free Update
```

- How long is your data around? (Chow et. al. '04)
  - Where in memory?
  - Maybe on disk?

Hard to Provide Sensitive Data Lifetime

Existing approaches fall short

- Shutdown the application?
- Reboot?
- Rely on application support?
- Memory scrubbing? (Chow et al '05: Data shredding)
- Change user behavior? (Borders et al '09: Capsules)
- Time-based data access control? (Perlman '05)

Goal: Guaranteed Data Lifetime

- Guarantee: Data indicated as sensitive is not retrievable from system beyond specified time limit

- Requirements
  - No application support

  - Non-disruptive : shouldn’t crash, interrupt your normal workflow

- Contribution: Promising start, much further to go

Observation: State Equivalence

- For any program state computed from sensitive data, there usually exists an equivalent state not derived from the sensitive data

- Example:
  - You get a sensitive email, read it, and then send and read some other emails
  - Equivalent State: Send and read other emails

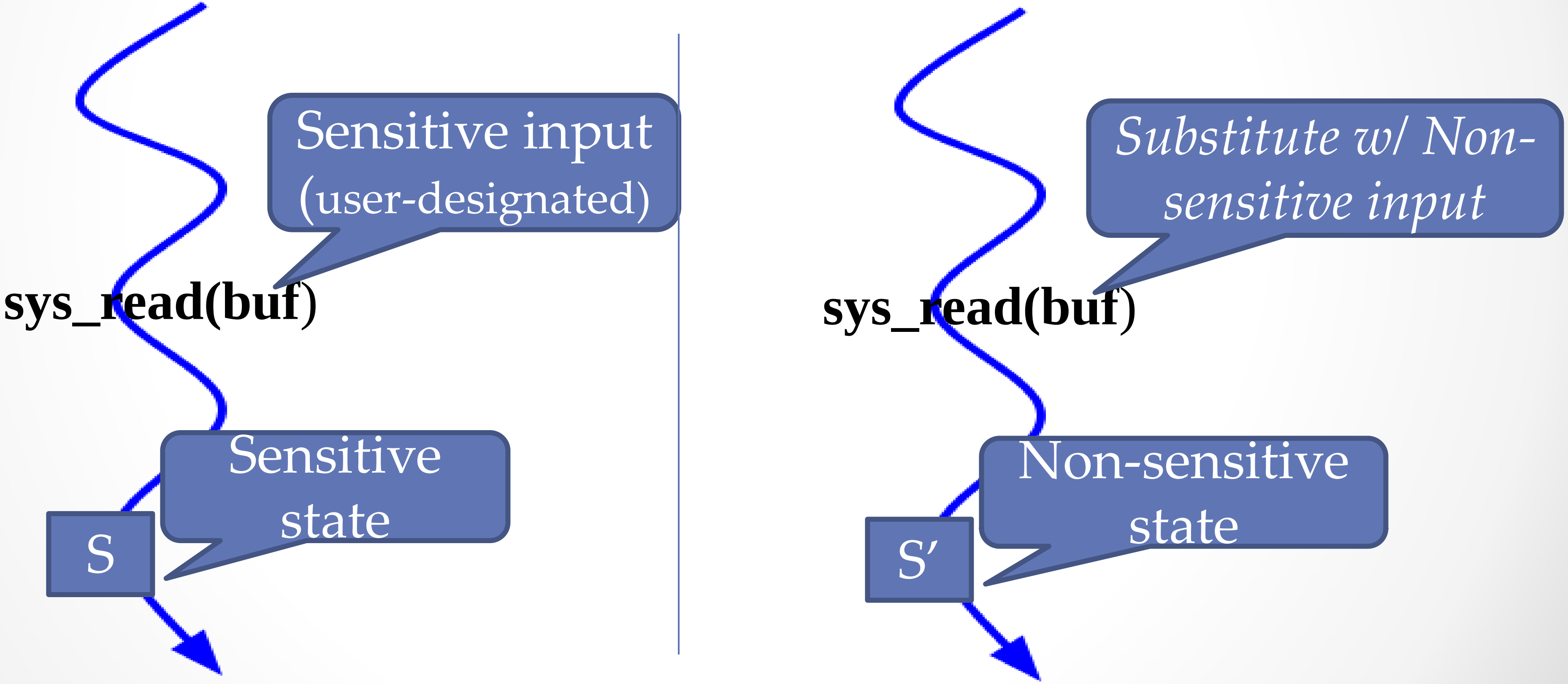
Approach: State Reincarnation

- Replace current sensitive state with equivalent non-sensitive state

- Challenge: How do we derive equivalent non-sensitive state?

Deriving an Equivalent State

- Key idea: deterministic replay with perturbed input


Sensitive input (user-designated)
sys_read(buf)
Sensitive state
S
Substitute w/ Non-sensitive input
sys_read(buf)
Non-sensitive state
S'

1. Original execution (record all inputs)

2. Replay execution (replace sensitive inputs)

Challenges

- Picking the sensitive-input replacements
- Completeness: Eliminating all sensitive data
- Overhead: Run-time cost

Picking sensitive-input replacements

- Given sensitive input I, and subsequent input I1, I2, we compute I' which leads to same execution path
    - Using tainting and constraint solving (Altekar '09)
- Replay with I'
- Hard-cases: Spell-checker, Hashing

Completeness

- Sensitive data can linger in various areas (OS buffers); how can we remove all of it?

- Technique: Implement perturbed replay in VM

- Need to trust VM not to retain data

Overhead

- We implemented recording at user-level
- Slowdown: ~1.2X on bash

Conclusion

- Contributions:
  - Guaranteed Lifetime Property
  - State Reincarnation

- Future work:
  - Picking right inputs for replay
  - Measuring overhead for consistent substitution