Exploring the Unchartered Space of Container Registry Typosquatting

Guannan Liu, Xing Gao, Haining Wang, Kun Sun

VIRGINIA TECH

UNIVERSITY OF DELAWARE

GEORGE MASON UNIVERSITY

usenix
THE ADVANCED COMPUTING SYSTEMS ASSOCIATION

Container Registry

- Container registries allow developers to publish, maintain, and manage images
- Public registry (Docker Hub, Quay.io)
  - Free unlimited storage for publicly accessible images
  - Free to download without authentication
- Private registry (Google, Amazon, IBM, etc.)
  - Image modification and download are authorized by the account owner
  - Charged based on storage and network usage
  - Image can be made public

Introduction

Overview

Exploitation

Mitigation

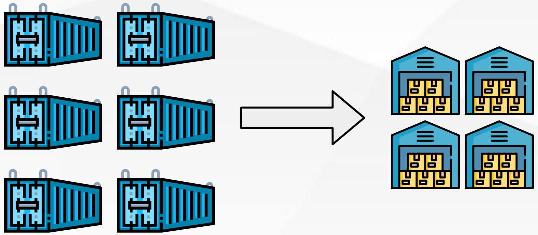
Conclusion

Container Image FQID

- Fully Qualified Image Identification
  - Used to uniquely identify an image
  - Distinguish images among many registries, usernames, and image names

registry_name/username/image_name

- Downloading an image
  - Docker Command-Line Interface (CLI): *$ docker pull*
  - Dockerfile: FROM statement
  - Users manually type FQID of the desired image



Introduction

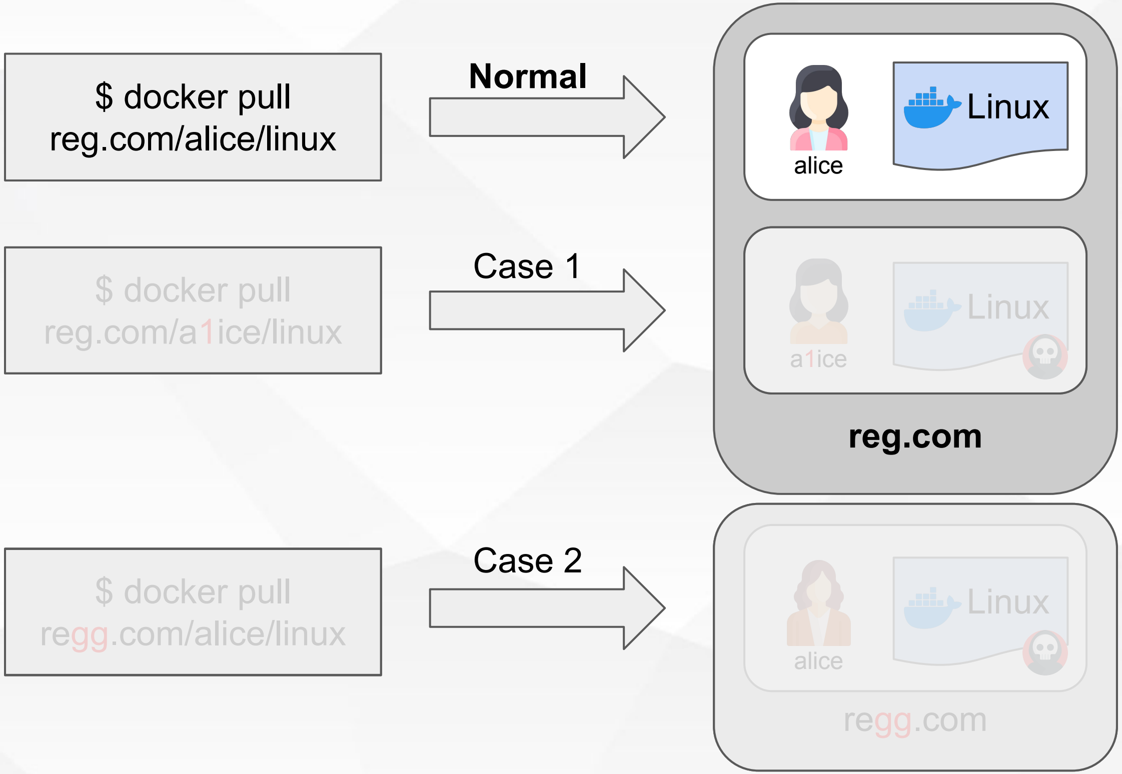
Overview

Exploitation

Mitigation

Conclusion

Container Registry Typosquatting



```
$ docker pull
reg.com/alice/linux
```

Normal

```
$ docker pull
reg.com/a1ice/linux
```

Case 1

```
$ docker pull
regg.com/alice/linux
```

Case 2

Introduction

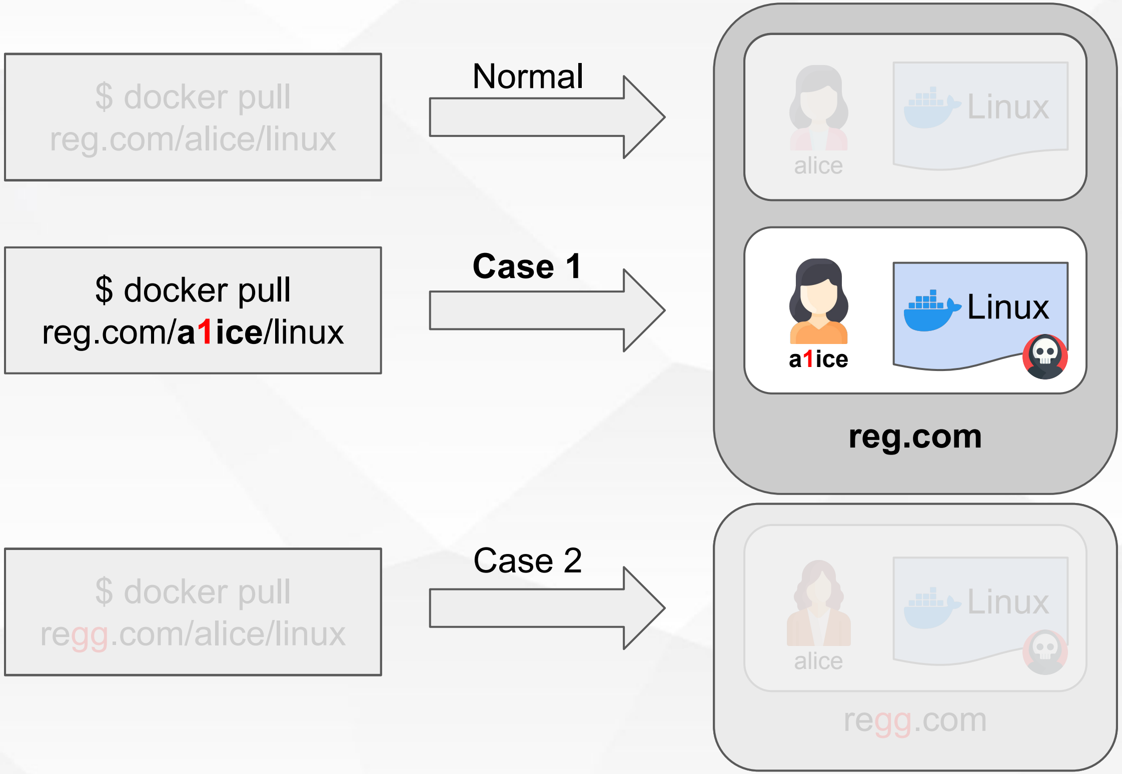
Overview

Exploitation

Mitigation

Conclusion

Container Registry Typosquatting



```
$ docker pull
reg.com/alice/linux
```

```
$ docker pull
reg.com/a1ice/linux
```

```
$ docker pull
regg.com/alice/linux
```

Introduction

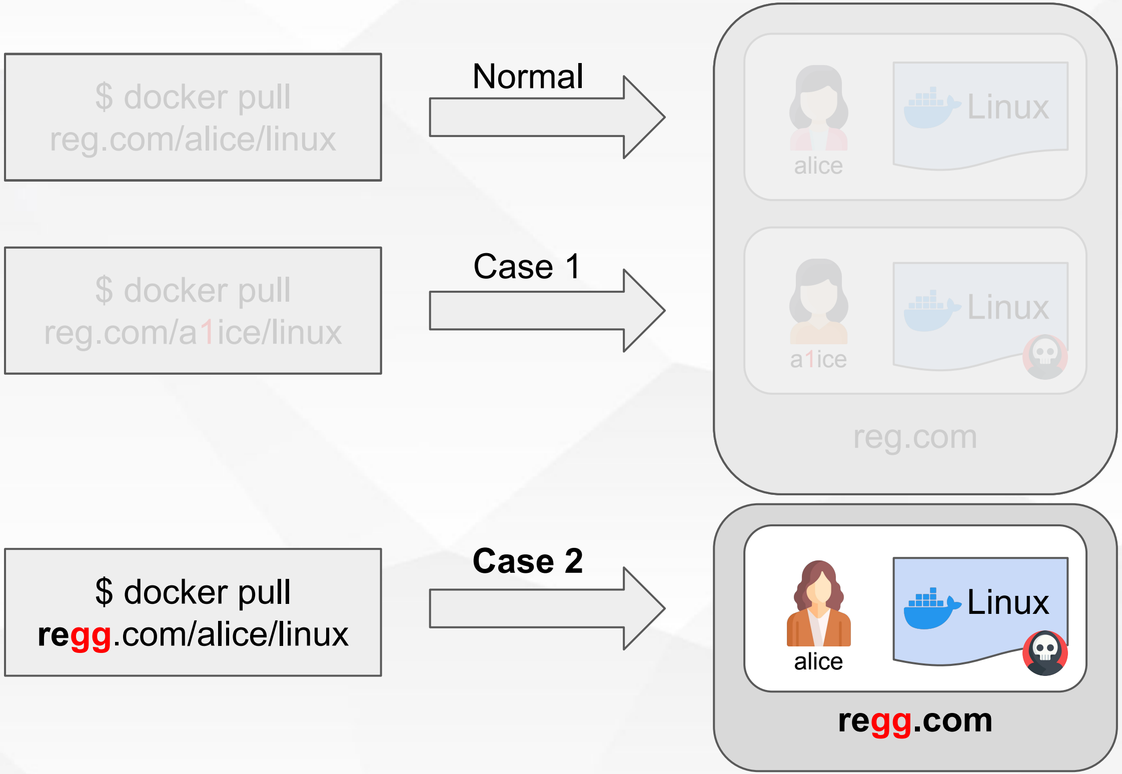
Overview

Exploitation

Mitigation

Conclusion

Container Registry Typosquatting



```
$ docker pull
reg.com/alice/linux
```

```
$ docker pull
reg.com/a1ice/linux
```

```
$ docker pull
regg.com/alice/linux
```

Introduction

Overview

Exploitation

Mitigation

Conclusion

Threat Model

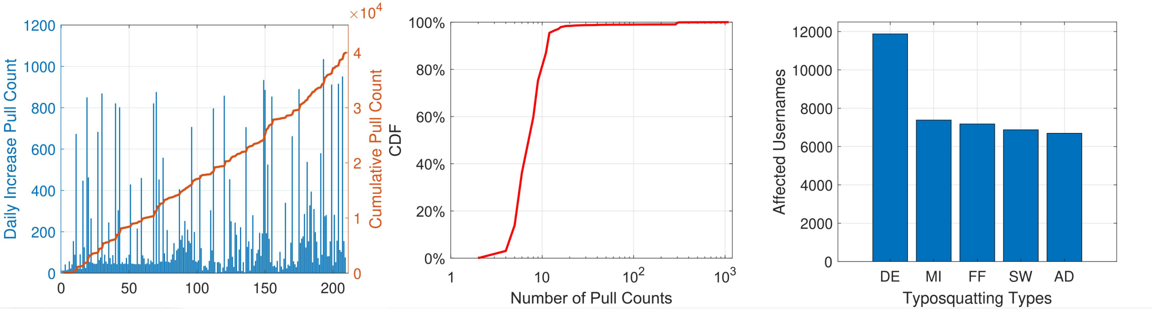
- Goal of attack
    - Generate multiple typosquatting FQIDs and bait users to pull images from a malicious repository
    - Distribute malicious container images by exploiting the potential typos made by container users
- Attack benefit
    - Generate financial profits
    - Obtain sensitive information of the victim
    - Harvest computing resources of the hosting server
    - Take control over the hosting server

Typosquatting in Public Registry

- Measurement on public registries
  - 416,087 and 21,409 container image repositories, with 246,080 and 6,475 unique usernames in Docker Hub and Quay.io
  - 75,312 typosquatting username pairs in Docker Hub and 119 pairs in Quay.io
- Proof-of-concept exploitation: 210 days
  - Target 10 existing Docker Hub usernames and generate 100 typosquatting usernames
  - Upload 4,787 images on Docker Hub with typosquatting FQIDs
  - Attract 40,009 total pull counts

Typosquatting in Public Registry

- Total pull count and daily increases
  - Linear increase trend: occurrence of mistyping FQID is random
  - Multiple daily spikes
- CDF distribution with respect to the number of pull counts
  - 37 most popular images attract 10,209 pull counts (largest: 1,094)
  - 80% images have pull counts of less than 10 but still attract 21,614 total pull counts
  - Popular images are more suitable for typosquatting attack, while less popular repositories might still be downloaded due to typing errors



Introduction

Overview

Exploitation

Mitigation

Conclusion

Typosquatting in Private Registry

- User-defined Project-ID (username) Typosquatting
  - Obtain 407 project-IDs from Alibaba, 158 from Azure, 407 from Google, and 584 from IBM
  - Randomly select 50 project-IDs from each registry for investigation
  - Generate full DL-1 typosquatting list for all project-IDs: 35,861, 32,629, 25,183, and 29,636 project-IDs for Alibaba, Azure, Google, and IBM
  - More than 90% of the DL-1 project-IDs are available for registration

- Randomly Generated Project-ID Typosquatting
  - Amazon randomly generates 12-digit client-ID as the project-ID
  - Register one piloting AWS account as benign ID and spawn 20,000 AWS accounts to attack
  - Only 1 DL-2 typosquatting project-ID is generated

Introduction

Overview

Exploitation

Mitigation

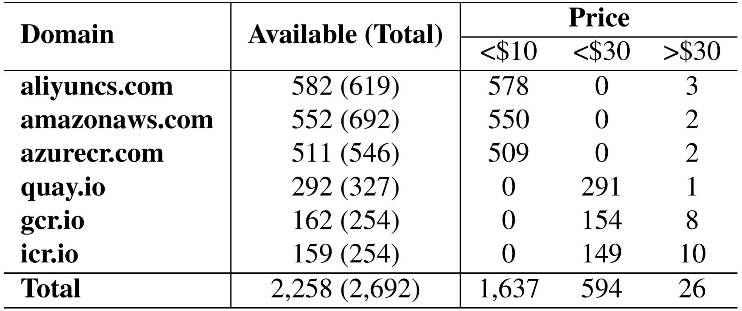
Conclusion

Typosquatting in Private Registry

- Proof-of-Concept Exploitation: 60-day
  - Target the official container images provided by Google
  - Select 10 images and generate 100 DL-1 typosquatting usernames
  - Record 62 pull counts for our uploaded images, with the highest download count of 14

Typosquatting Across Platforms

- Domain Typosquatting
  - Attackers self-host typosquatting container registries
  - Generate 2,692 DL-1 typosquatting domain names for the six container registries
  - 2,258 (83.9%) of them are available for purchase
  - 72.5% (1,637 out of 2,258) domains cost less than $10, and 26 domains have a purchasing price over $30



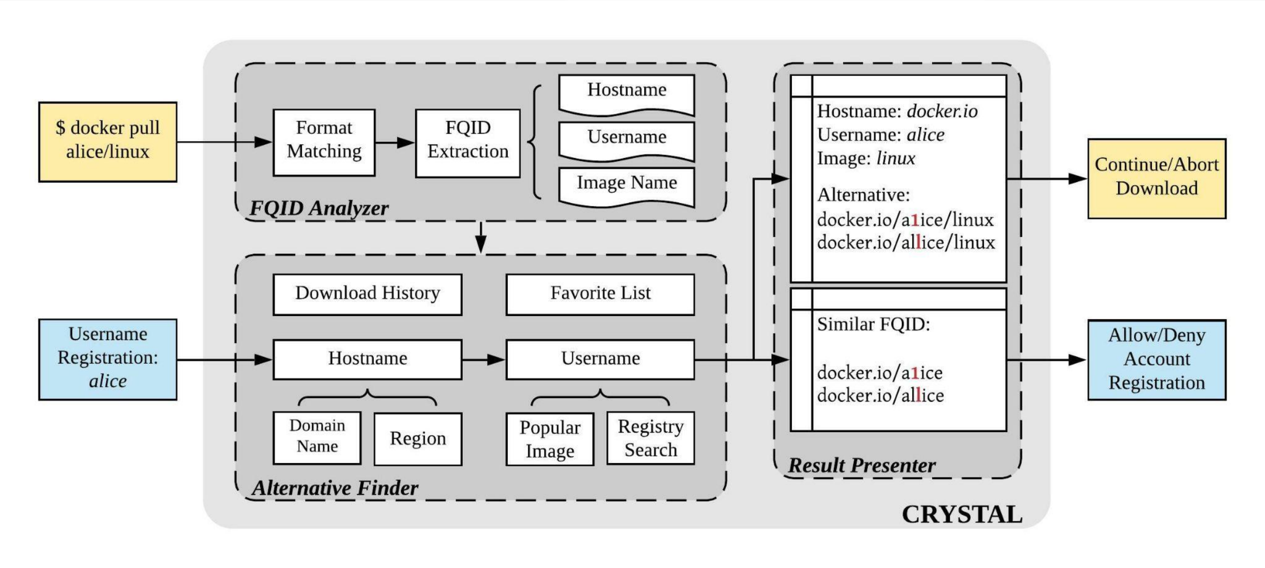
Domain	**Available (Total)**	**Price**		
		<$10	<$30	>$30
aliyuncs.com	582 (619)	578	0	3
amazonaws.com	552 (692)	550	0	2
azurecr.com	511 (546)	509	0	2
quay.io	292 (327)	0	291	1
gcr.io	162 (254)	0	154	8
icr.io	159 (254)	0	149	10
Total	2,258 (2,692)	1,637	594	26

Typosquatting Across Platforms

- Missing Hostname
  - By default, hostname can be omitted if the container image is hosted in Docker Hub
  - Users who forget to include a hostname in the Docker pull command might obtain an unwanted image from Docker Hub
  - Select 10 usernames from Quay.io and register them on Docker Hub
  - Record 93 pull counts in 30-day experiments, with the highest pull counts of 24

Mitigation

- CRYSTAL (Container RegistrY SquaTting ALarm)
- Used on user’s side to alert typing error, as well as on registry’s side to prevent typosquatting usernames



Introduction

Overview

Exploitation

Mitigation

Conclusion

Conclusion

- Systematic study on container registry typosquatting
    - Users indeed make typing mistakes when downloading docker images
    - A large amount of typosquatting usernames, project-IDs, and domain names are currently available for public registration
    - Pose realistic security threats to the container ecosystem
- Propose mitigation tool: CRYSTAL
    - Alert users about potential typing errors
    - Assist container registries to discover potential typosquatting FQIDs
    - Achieve a high detection accuracy of more than 97.5% with low overhead

Introduction

Overview

Exploitation

Mitigation

Conclusion

Questions?

guannanliu@vt.edu

Our paper in
USENIX Website

Guannan Liu’s
Personal Website

usenix
THE ADVANCED
COMPUTING SYSTEMS
ASSOCIATION