浙江大学网络系统安全与隐私实验室
NETWORK SYSTEM SECURITY & PRIVACY LAB

Unveiling Security Vulnerabilities in Git Large File Storage Protocol

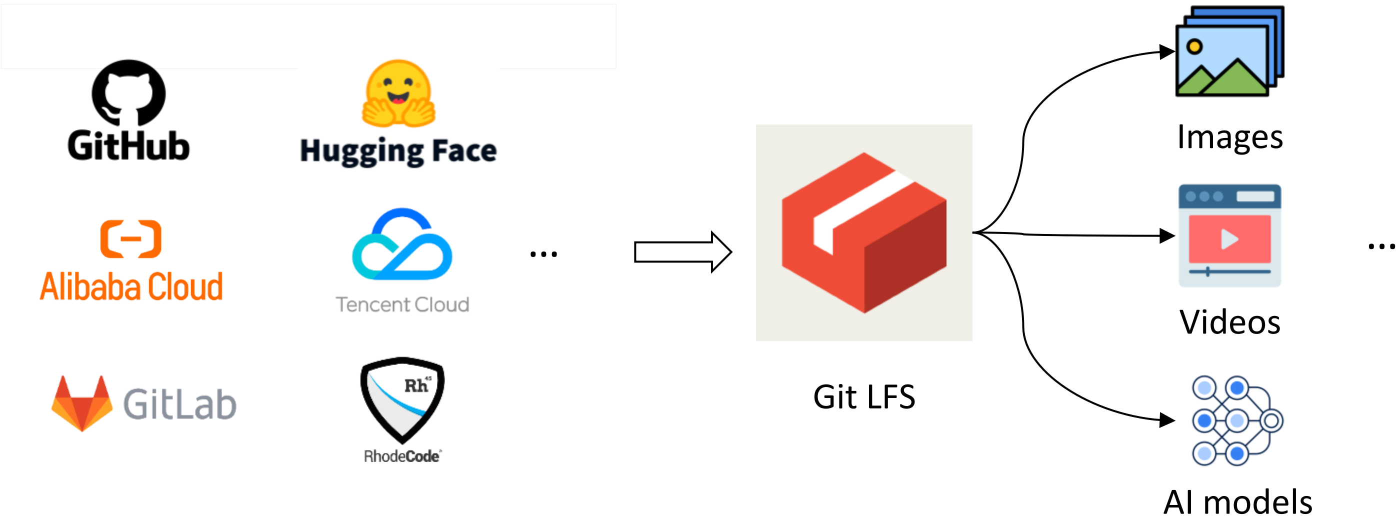
Yuan Chen, **Qinying Wang**, Yong Yang, Yuanchao Chen,
Yuwei Li, Shouling Ji

浙江大學
ZHEJIANG UNIVERSITY

国防科技大学
NATIONAL UNIVERSITY
OF DEFENSE TECHNOLOGY

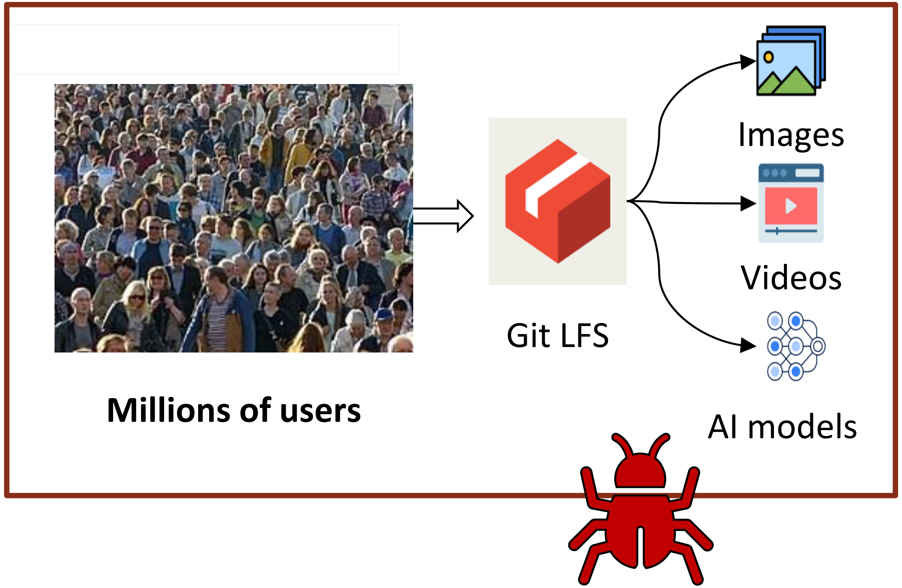
Motivation & Background

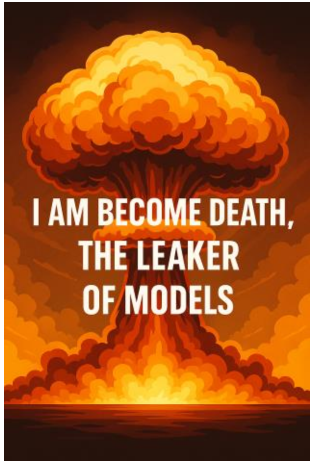
Git Large File Storage (LFS) is a **widely** used Git extension for managing large files and binaries: 23+ million LFS files on GitHub; 1+ million AI models on Hugging Face, …

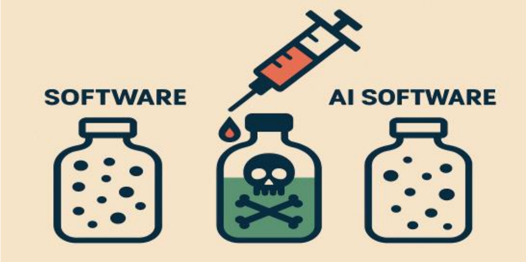


Motivation & Background

A compromise of LFS can result in serious consequences, including **sensitive data leakage** and **supply chain poisoning**.







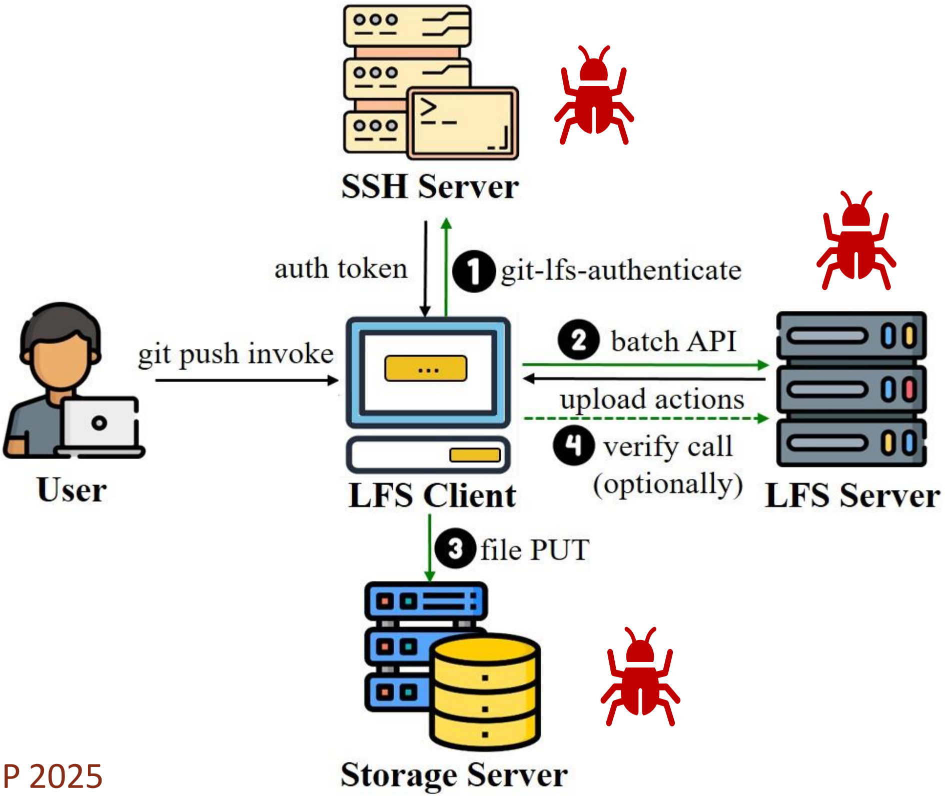
Motivation & Background

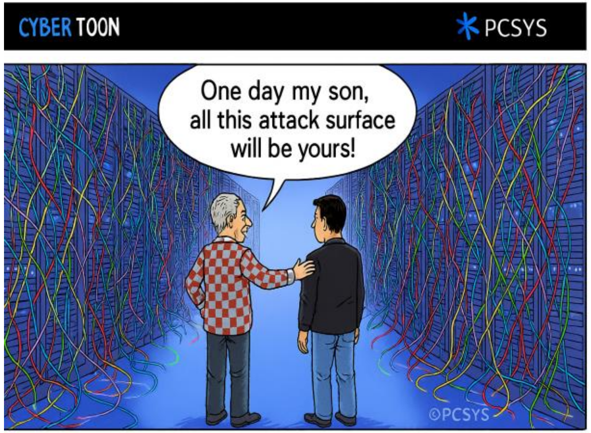
A compromise of LFS can result in serious consequences, including **sensitive data leakage** and **supply chain poisoning**.

The security and risks of Git LFS remain largely unexplored!

Challenge 1: Protocol Complexity

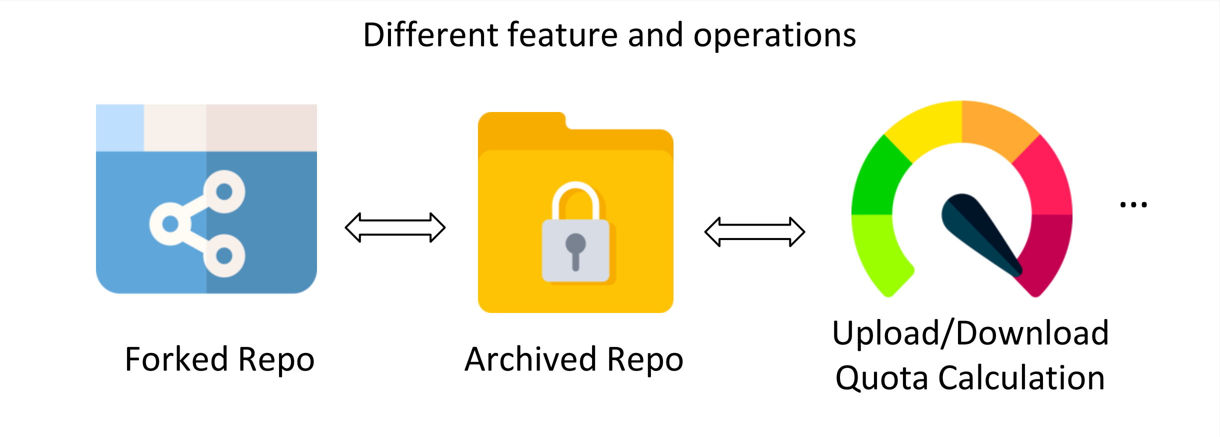
The LFS protocol interactions are **complex**, introducing new attack surfaces.





Challenge 2: Compositional Intricacy

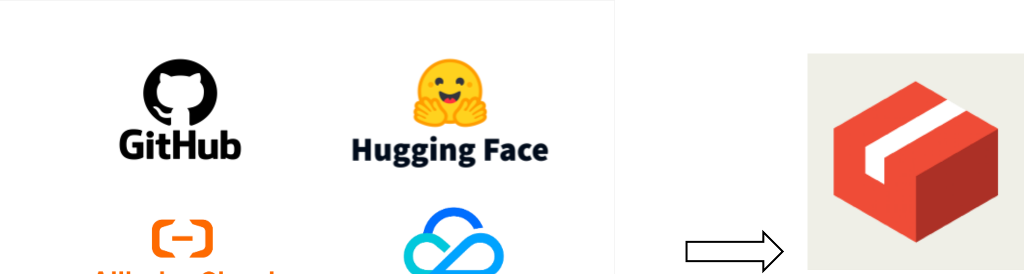
The **interplay** between Git LFS and auxiliary features introduces subtle and often over looked security risks.


Different feature and operations
Forked Repo
Archived Repo
Upload/Download Quota Calculation
...

LFS Upload -> Quota Escape !

☁ Challenge 3: Infrastructure Heterogeneity

The **varied implementations** significantly impede systematic vulnerability detection across different platforms.



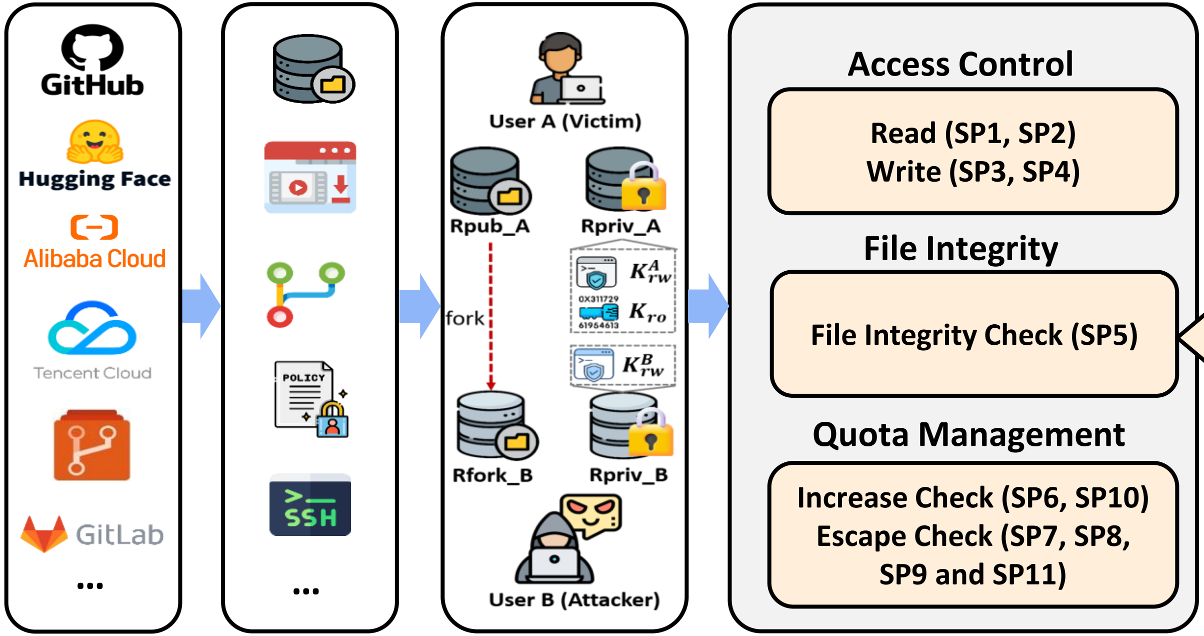
Git LFS

- Different cloud storage
- Different API call enforcement
- Different quota enforcement

I AM NOT SAYING IT IS A CUSTOM IMPLEMENTATION
BUT IT IS A CUSTOM IMPLEMENTATION
imgflip.com

Solution

We adopt a **feature-based, property-driven** approach.



Git Platforms

Supported Features

Preparation

Access Control

Read (SP1, SP2)
Write (SP3, SP4)

File Integrity

File Integrity Check (SP5)

Quota Management

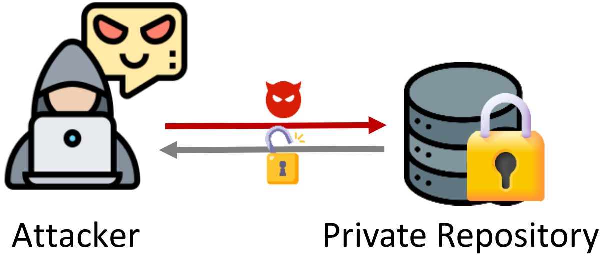
Increase Check (SP6, SP10)
Escape Check (SP7, SP8, SP9 and SP11)

Property Checking

SP5: Uploaded LFS objects must be verified to ensure their content matches the claimed SHA-256 hash, to prevent the upload of tampered files.

New Attacks

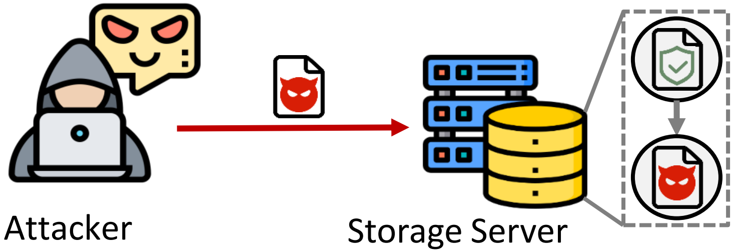
We further identified **four** new attacks vectors.



Attacker

Private Repository

1. LFS File Leakage



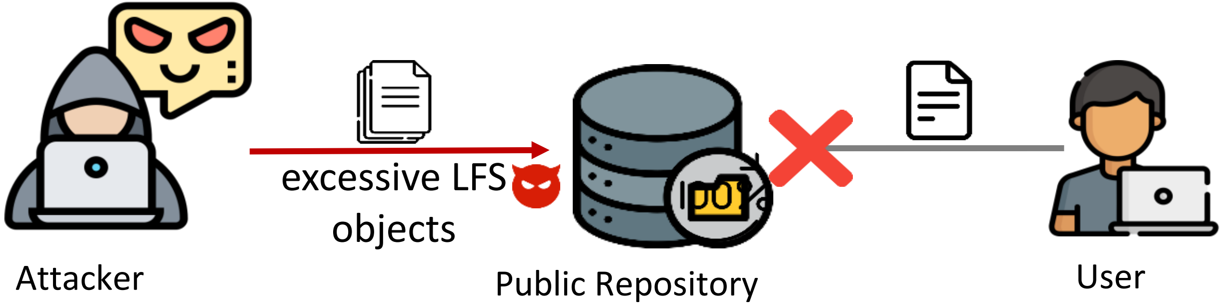
Attacker

Storage Server

2. LFS File Replacement

New Attacks

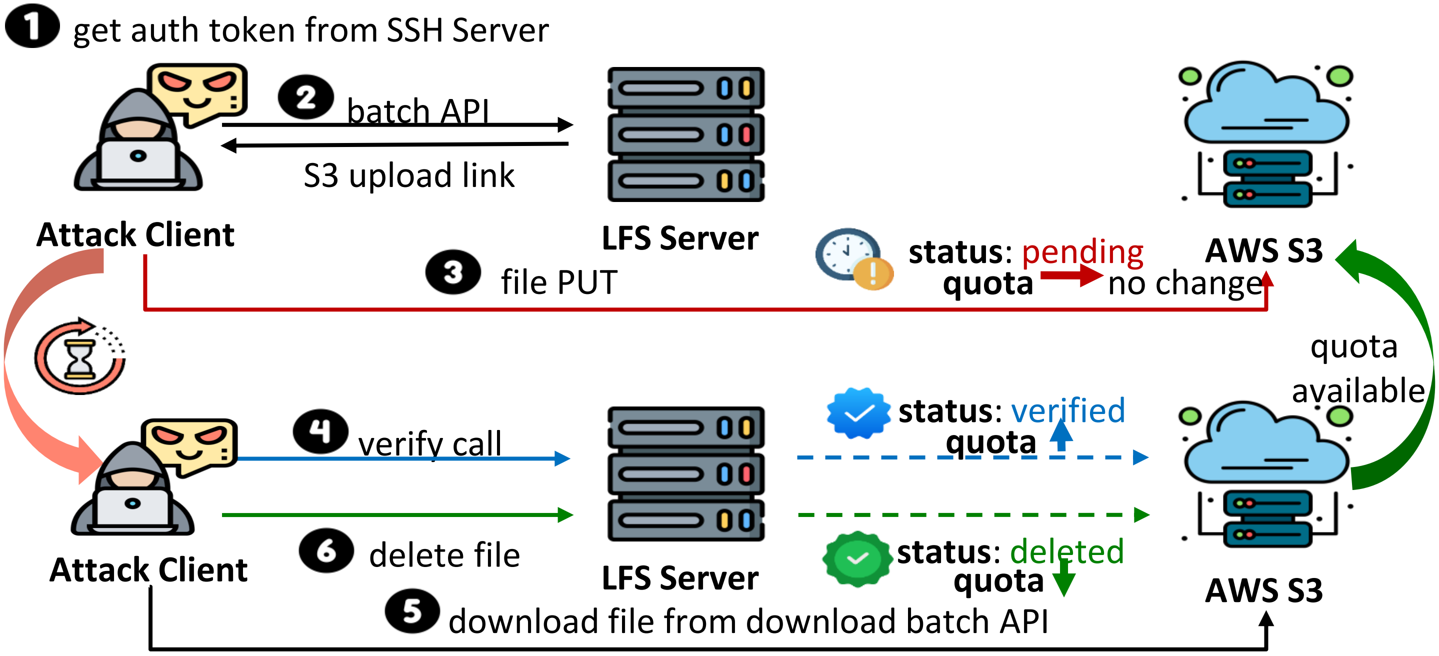
We further identified **four** new attacks vectors.



3. Quota-Based Denial of Service Attack

New Attacks

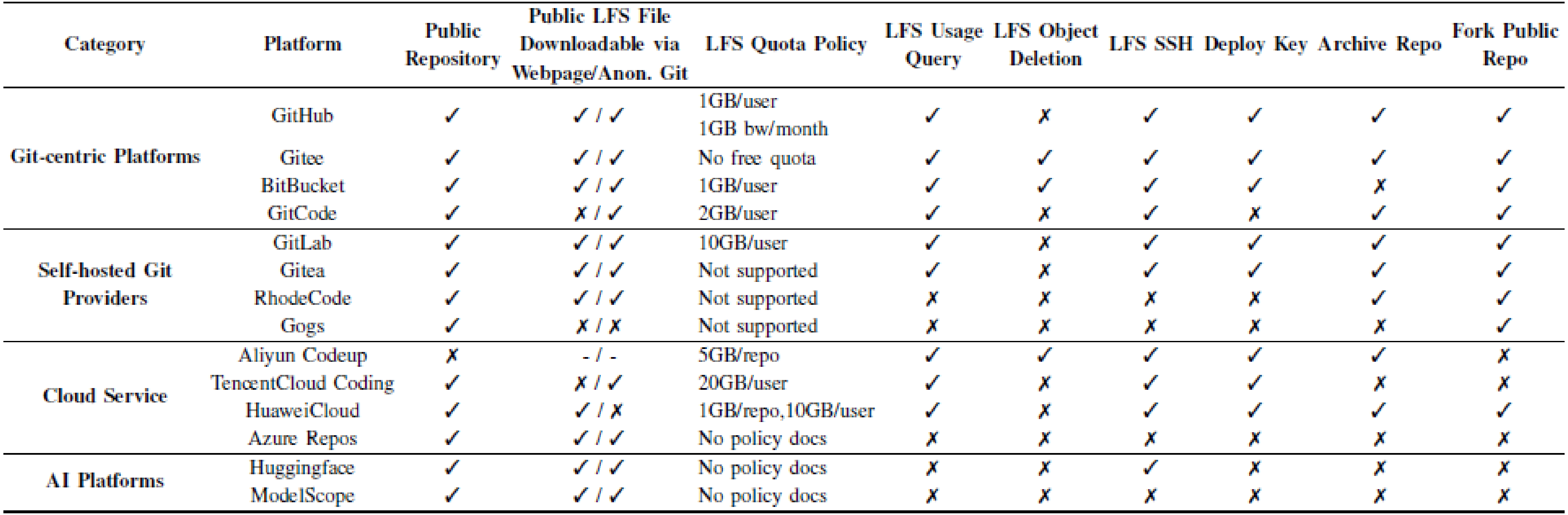
We further identified **four** new attacks vectors.



4. Quota Escape Example: Delay Attack

Real-World Evaluation

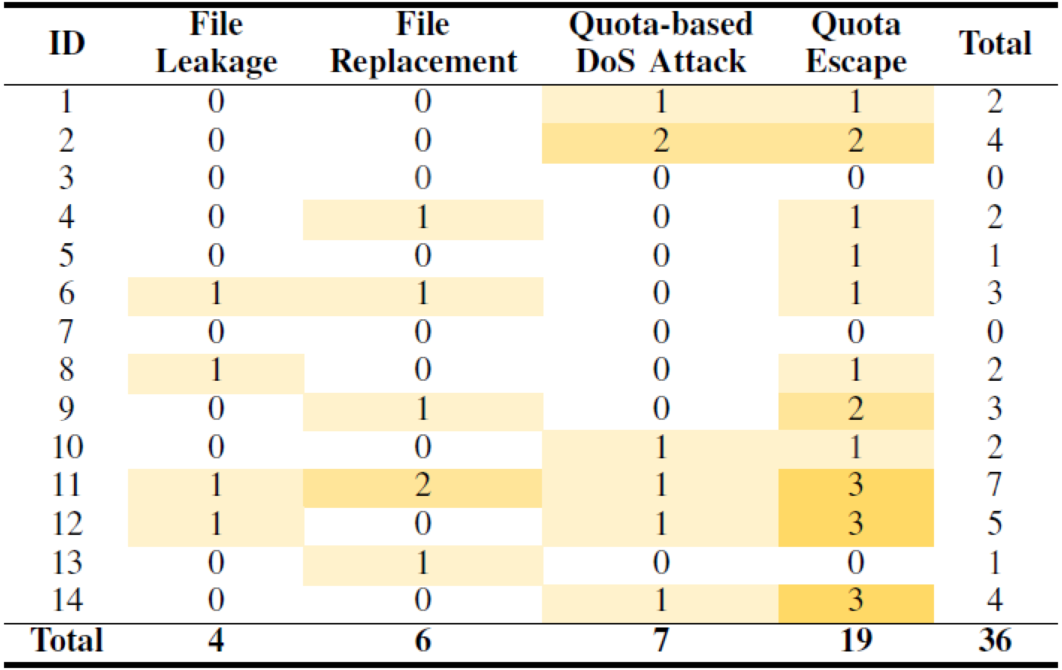
We evaluate our framework on real world **14** Git platforms spanning **four categories**.



Category	Platform	Public Repository	Public LFS File Downloadable via Webpage/Anon. Git	LFS Quota Policy	LFS Usage Query	LFS Object Deletion	LFS SSH	Deploy Key	Archive Repo	Fork Public Repo
Git-centric Platforms	GitHub	✓	✓ / ✓	1GB/user 1GB bw/month	✓	✗	✓	✓	✓	✓
	Gitee	✓	✓ / ✓	No free quota	✓	✓	✓	✓	✓	✓
	BitBucket	✓	✓ / ✓	1GB/user	✓	✓	✓	✓	✗	✓
	GitCode	✓	✗ / ✓	2GB/user	✓	✗	✓	✗	✓	✓
Self-hosted Git Providers	GitLab	✓	✓ / ✓	10GB/user	✓	✗	✓	✓	✓	✓
	Gitea	✓	✓ / ✓	Not supported	✓	✗	✓	✓	✓	✓
	RhodeCode	✓	✓ / ✓	Not supported	✗	✗	✗	✗	✓	✓
	Gogs	✓	✗ / ✗	Not supported	✗	✗	✗	✗	✗	✓
Cloud Service	Aliyun Codeup	✗	- / -	5GB/repo	✓	✓	✓	✓	✓	✗
	TencentCloud Coding	✓	✗ / ✓	20GB/user	✓	✗	✓	✓	✗	✗
	HuaweiCloud	✓	✓ / ✗	1GB/repo,10GB/user	✓	✗	✓	✓	✓	✓
	Azure Repos	✓	✓ / ✓	No policy docs	✗	✗	✗	✗	✗	✗
AI Platforms	Huggingface	✓	✓ / ✓	No policy docs	✗	✗	✓	✗	✗	✗
	ModelScope	✓	✓ / ✓	No policy docs	✗	✗	✗	✗	✗	✗

Real-World Evaluation

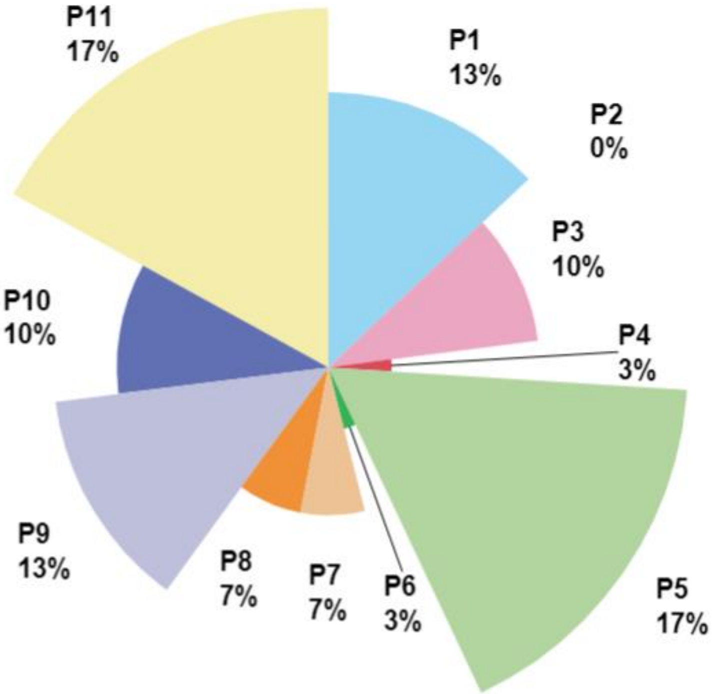
Our analysis uncovered **36 vulnerabilities** across **14 Git platforms**, impacting a broad user base and confirmed by the affected vendors.



ID	File Leakage	File Replacement	Quota-based DoS Attack	Quota Escape	Total
1	0	0	1	1	2
2	0	0	2	2	4
3	0	0	0	0	0
4	0	1	0	1	2
5	0	0	0	1	1
6	1	1	0	1	3
7	0	0	0	0	0
8	1	0	0	1	2
9	0	1	0	2	3
10	0	0	1	1	2
11	1	2	1	3	7
12	1	0	1	3	5
13	0	1	0	0	1
14	0	0	1	3	4
Total	**4**	**6**	**7**	**19**	**36**

Real-World Evaluation

Our analysis revealed that nearly **all** security properties have corresponding violations, with SP5 and SP11 having the highest number of violations(five each).



Unveiling Security Vulnerabilities in Git Large File Storage Protocol

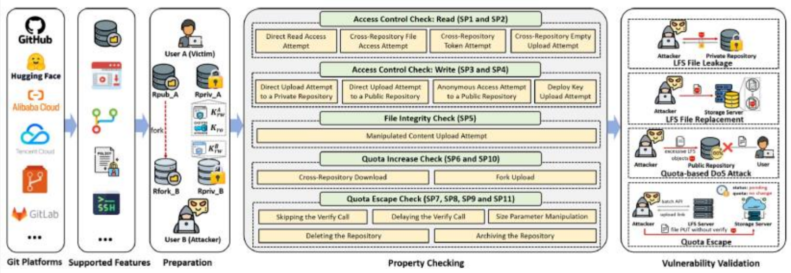
Four New Attacks

Challenge 1: Protocol Complexity
Challenge 2: Compositional Intricacy
Challenge 3: Infrastructure Heterogeneity

File Leakage, File Replacement, Quota-based DoS, Quota Escape

Semi-automated Testing Framework

A feature-based, property-driven philosophy.



Real World Impact

GitHub
Hugging Face
Alibaba Cloud
Tencent Cloud
...
GitLab
RhodeCode

- 14 major platforms
- 36 unknow vulnerabilities
- $1800 bug bounty

We are actively collaborating with GitHub to enhance the LFS protocol!

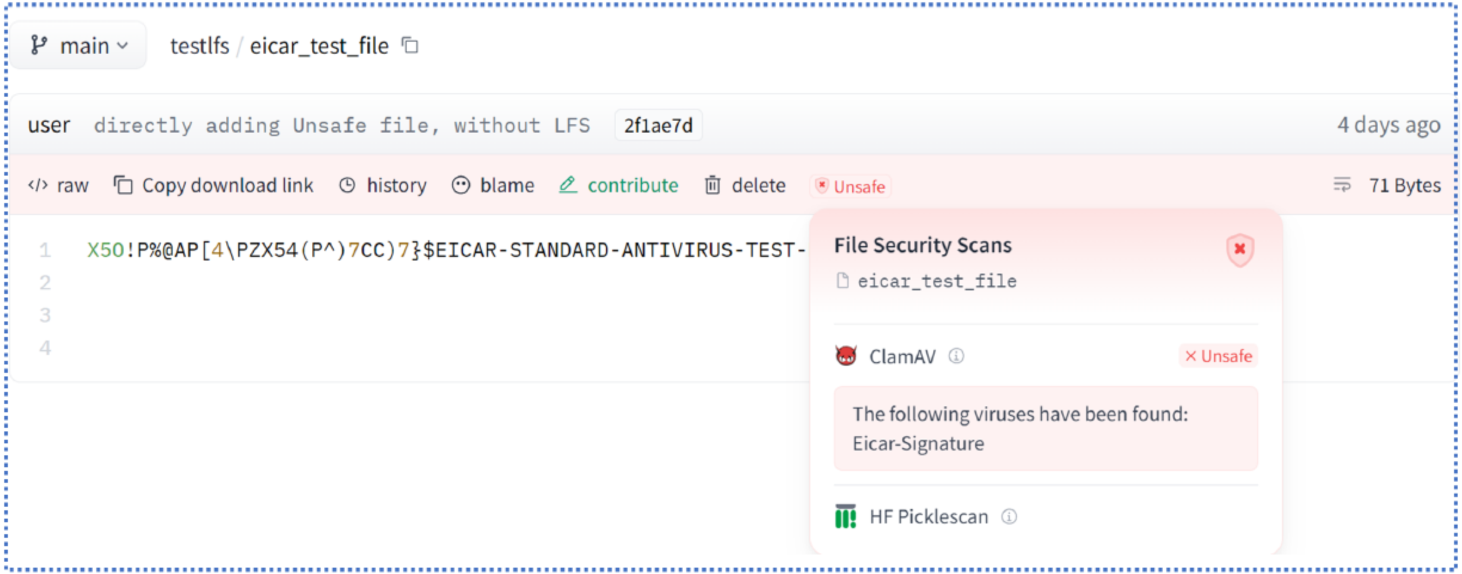
浙江大學
ZHEJIANG UNIVERSITY

Email: chenyuan@zju.edu.cn
wqybbh@gmail.com
yangyong202@zju.edu.cn

国防科技大学
NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY

Case Study

We demonstrate the critical impact of the file overwrite vulnerability.



user directly adding Unsafe file, without LFS 2f1ae7d 4 days ago

```
X50!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-
```

File Security Scans

eicar_test_file

ClamAV ✕ Unsafe

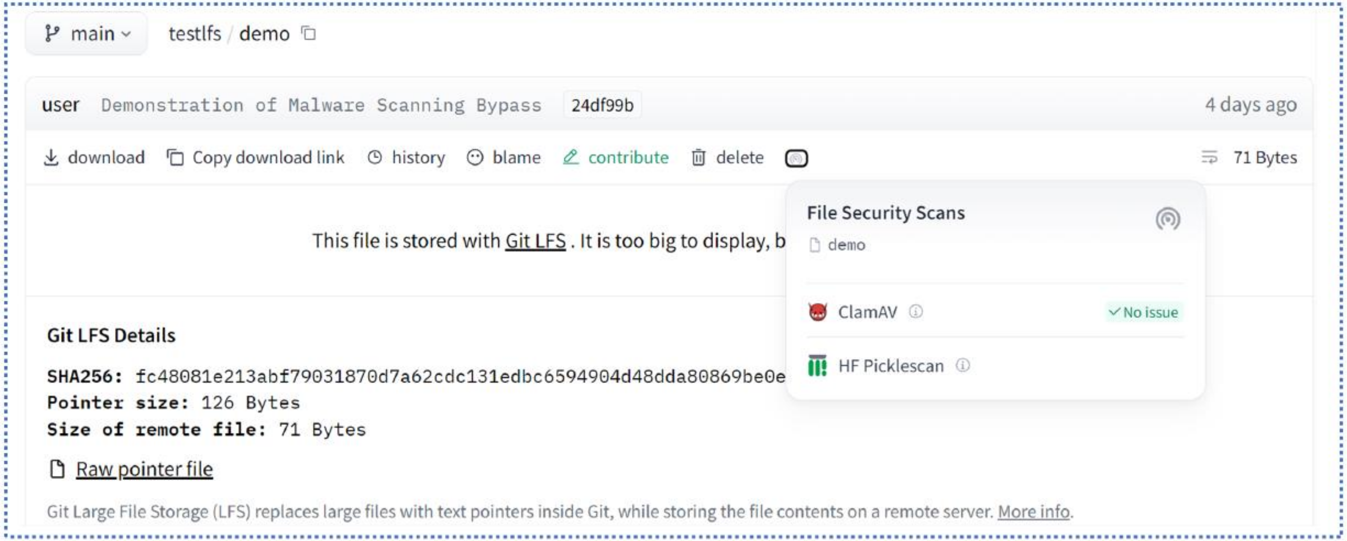
The following viruses have been found:
Eicar-Signature

HF Picklescan

Case Study

We demonstrate the critical impact of the file overwrite vulnerability.

Upload link:
https://<redacted_s3_domain>/repos/<repoid_prefix>/<repoid>/<sha256>
?X-Amz-Content-Sha256=**UNSIGNED-PAYLOAD**&X-Amz-Expires=**900**&X-Amz-Signature=<signature>&...



main testlfs / demo

user Demonstration of Malware Scanning Bypass 24df99b 4 days ago

download Copy download link history blame contribute delete 71 Bytes

This file is stored with Git LFS . It is too big to display, b

File Security Scans

demo

ClamAV ✓ No issue

HF Picklescan

Git LFS Details

SHA256: fc48081e213abf79031870d7a62cdc131edbc6594904d48dda80869be0e
Pointer size: 126 Bytes
Size of remote file: 71 Bytes

Raw pointer file

Git Large File Storage (LFS) replaces large files with text pointers inside Git, while storing the file contents on a remote server. More info.