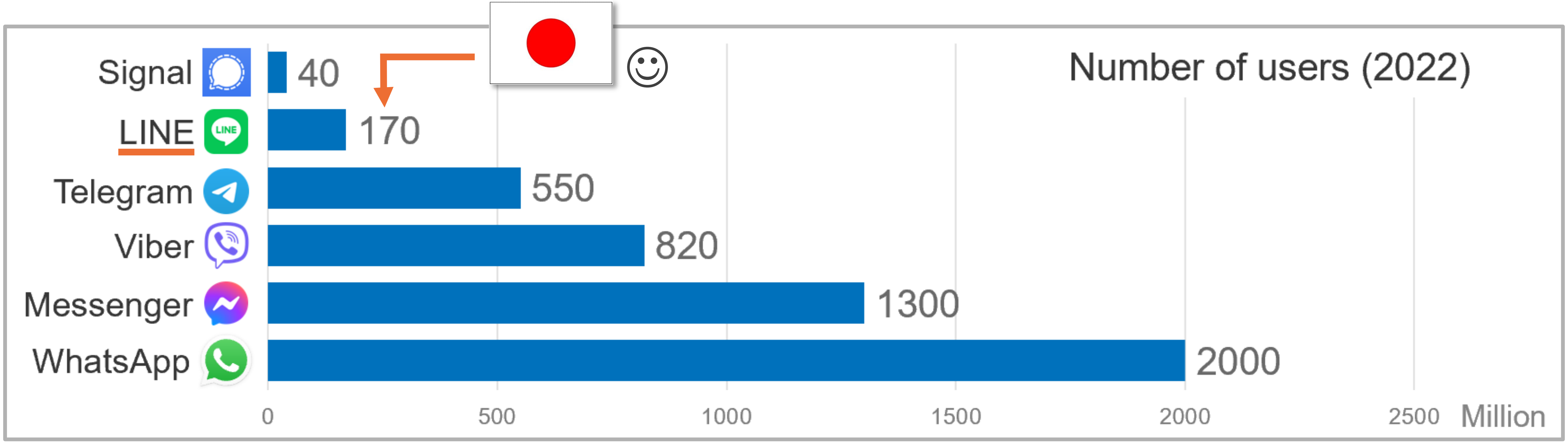
(Post-Quantum) Metadata Protection for MLS and Its Variants

Keitaro Hashimoto (1,3) **Shuichi Katsumata** (2,3) Thomas Prest (2)

Tokyo Tech (1) PQSHIELD (2) AIST (3)

Messaging Apps are Used Worldwide



https://www.businessofapps.com/data/messaging-app-market/

Unique Security Threats

□ Messaging apps are mostly run on our smartphones, which are deeply tied to our daily lives.

⇒ Contents and metadata tells a LOT about us.

□ Longer conversations compared to TLS-based communications, e.g., web browsing.

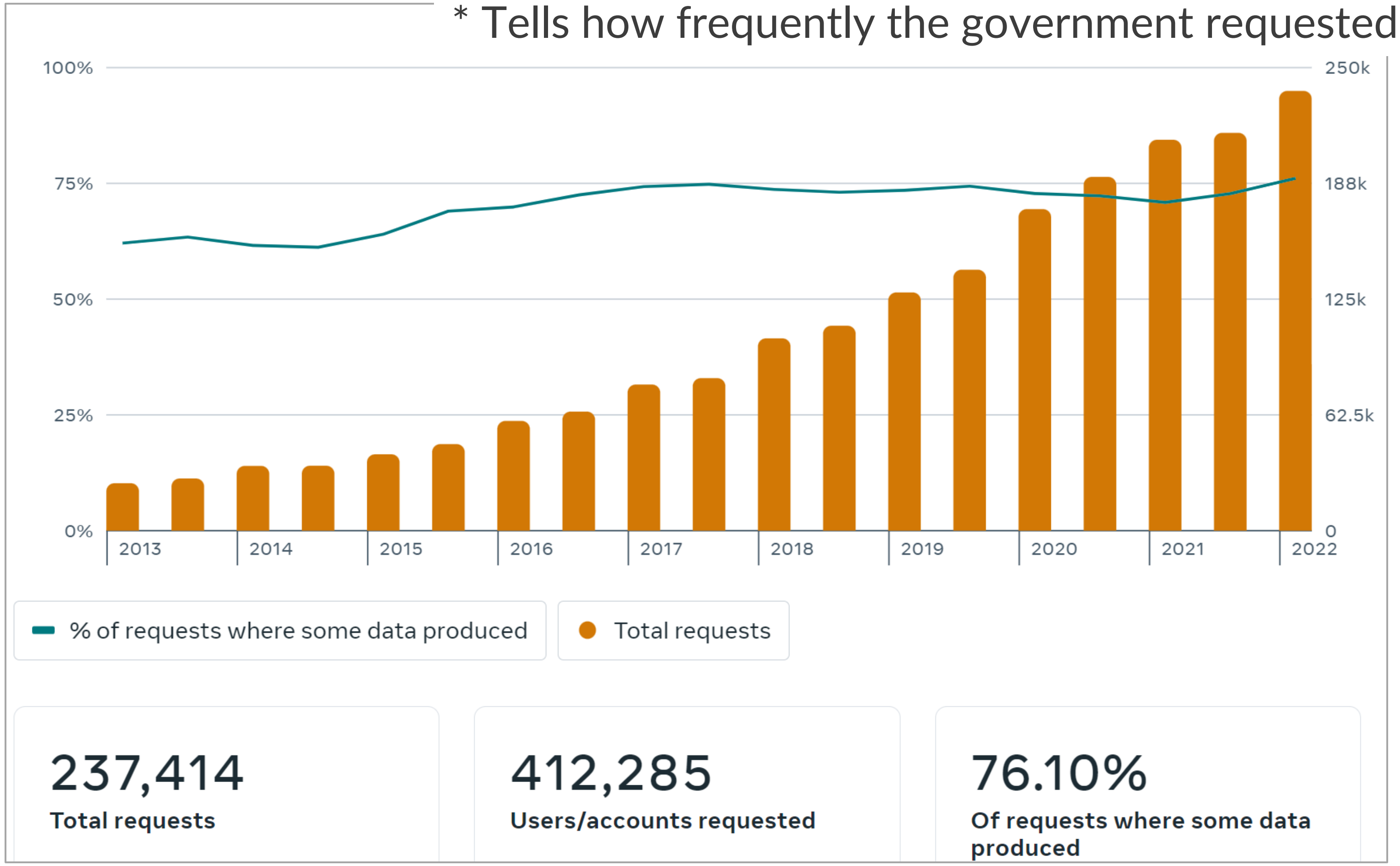
⇒ Higher risk of security breach
+ insecure period persists longer.

free background photos from https://pngtree.com/freebackground/dark-blue-technology-background-with-digital-network-lines-and-illustration-of-smartphone-and-message-balloon_1238300.html?sol=downref&id=bef

Your Personal Data can be Attack

Meta’s transparency report from Jan-Jun 2022.

* Tells how frequently the government requested data in a given period.



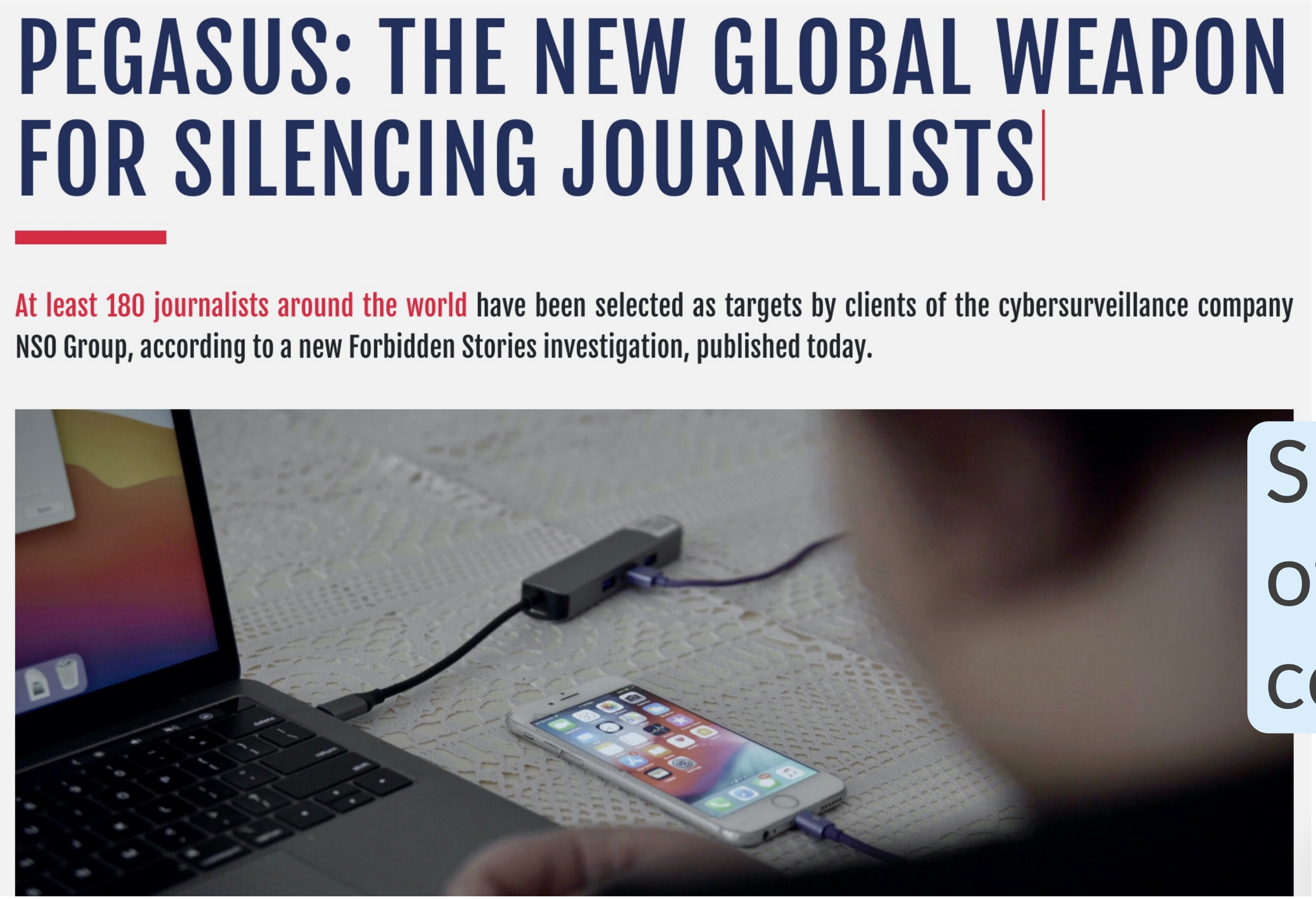
237,414
Total requests

412,285
Users/accounts requested

76.10%
Of requests where some data produced

https://transparency.fb.com/data/government-data-requests/

Legal Means are Not the End of the Story



PEGASUS: THE NEW GLOBAL WEAPON FOR SILENCING JOURNALISTS

At least 180 journalists around the world have been selected as targets by clients of the cybersurveillance company NSO Group, according to a new Forbidden Stories investigation, published today.

Spyware are sold off-the-shelf by companies and hackers.

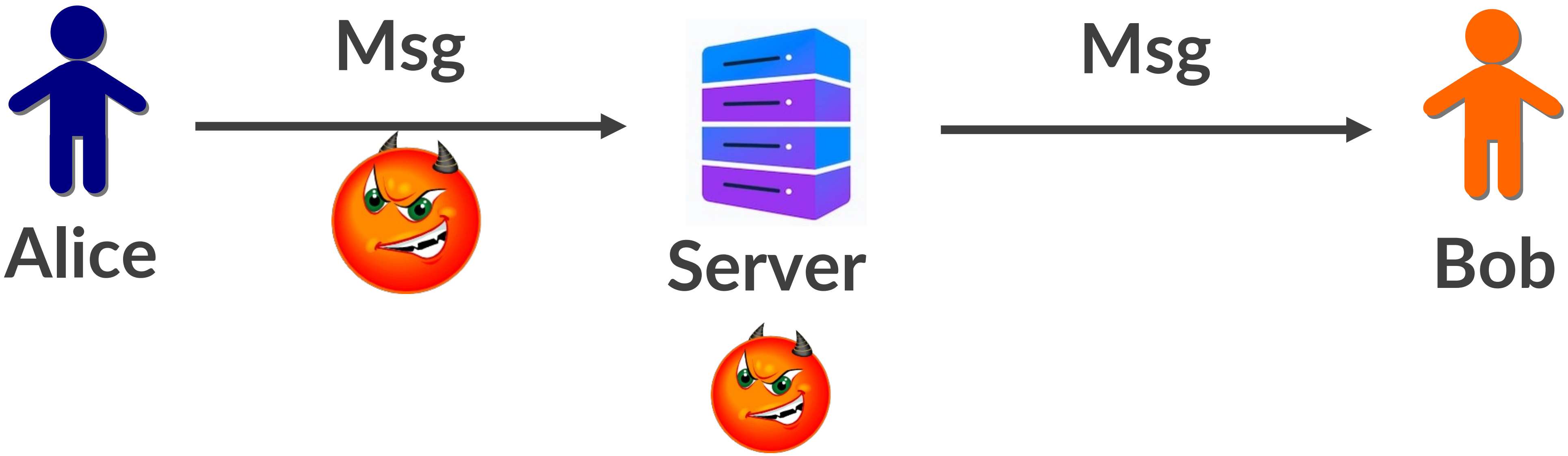
https://forbiddenstories.org/pegasus-the-new-global-weapon-for-silencing-journalists/

Today’s Talk: Scalable Secure Group Messaging

1. A systematic view on “security”
2. Comparing Signal and MLS
3. Making MLS metadata hiding

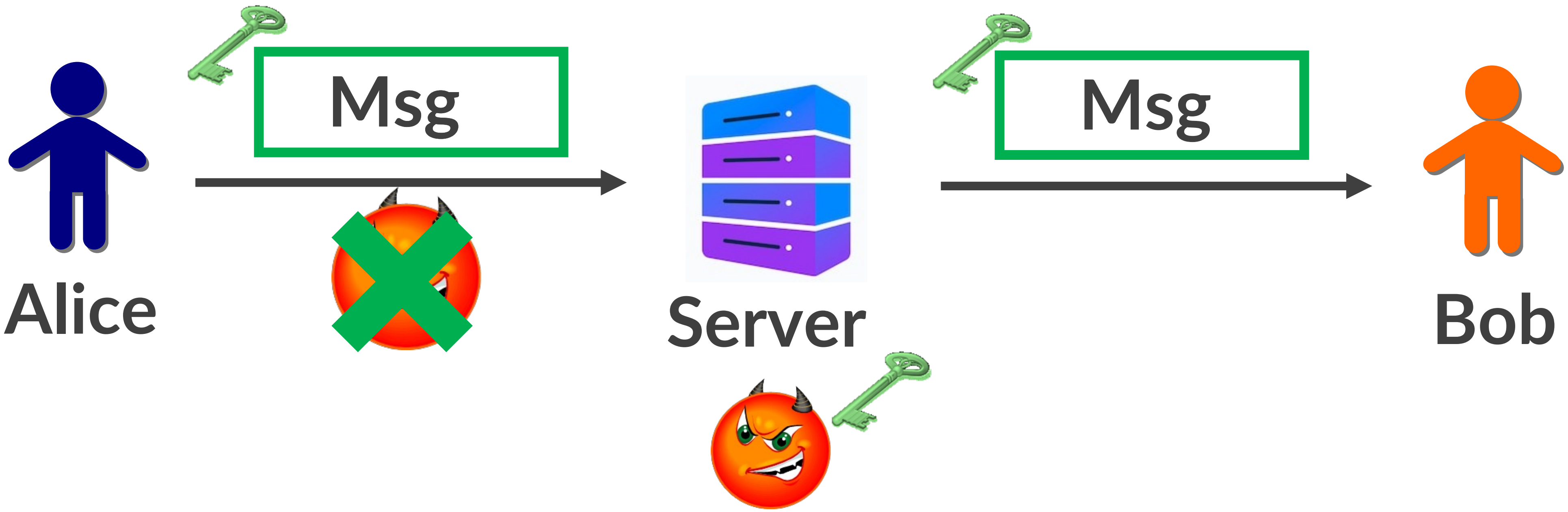
Taking a Closer Look at Messaging Protocols

Alice wants to "securely" send a message to Bob.



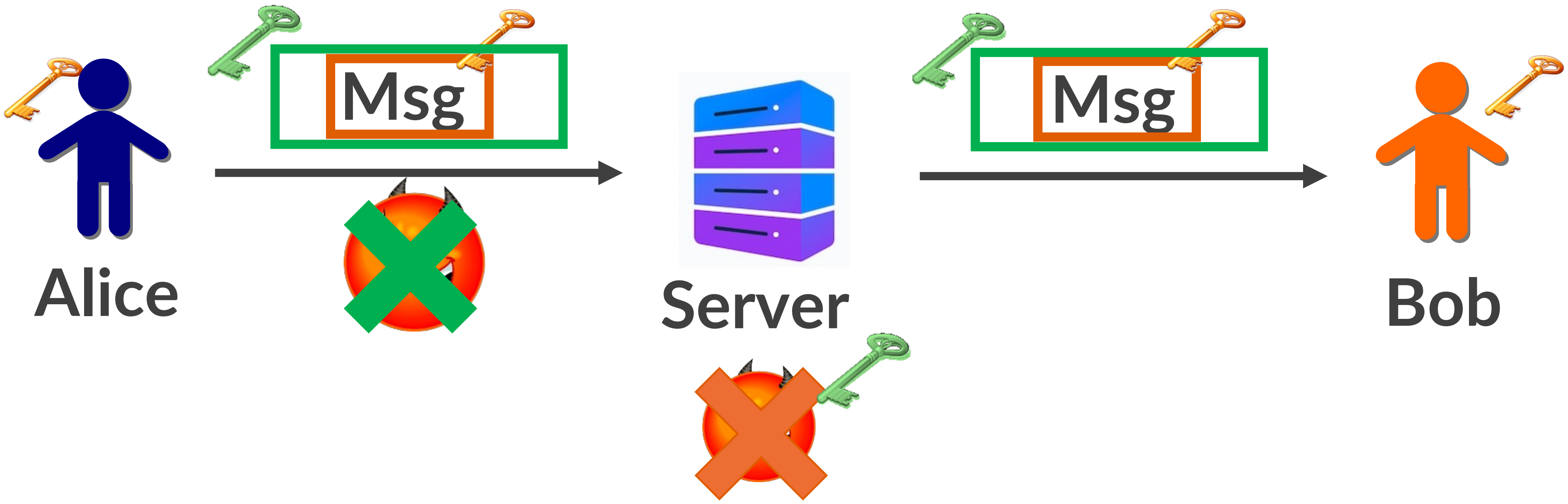
Removing Outsider Threats

E2EE between user and server (e.g., TLS/Noise)



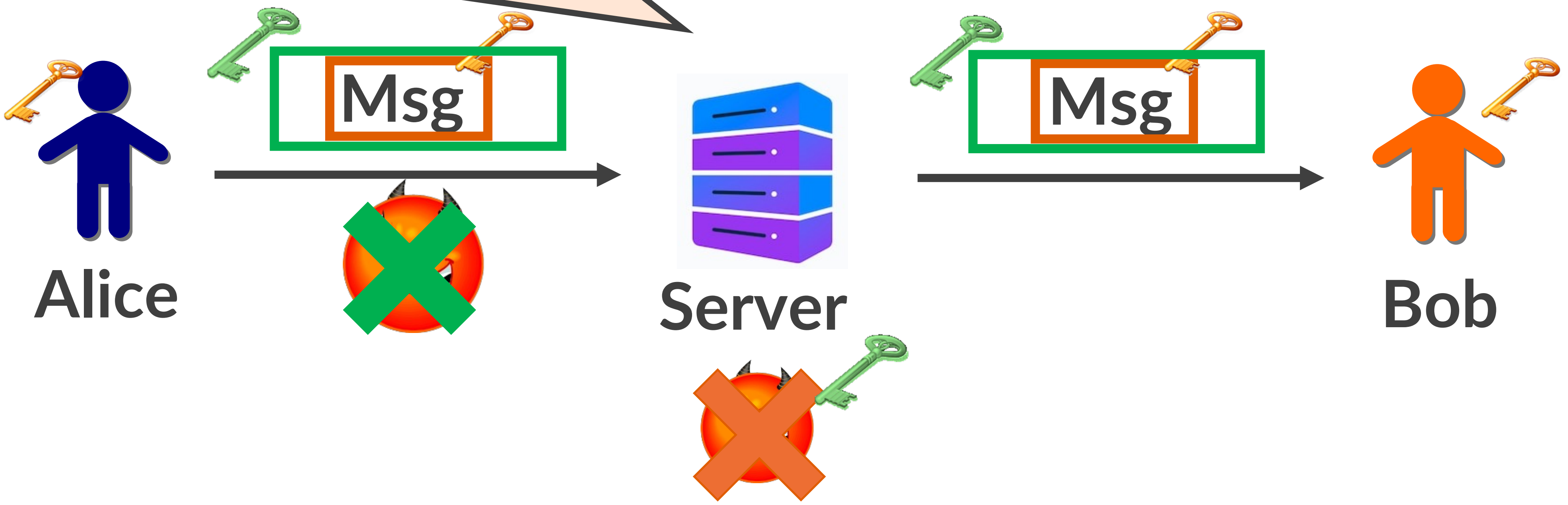
Removing Insider Threats

E2EE between users Alice and Bob



Removing Insider Threats

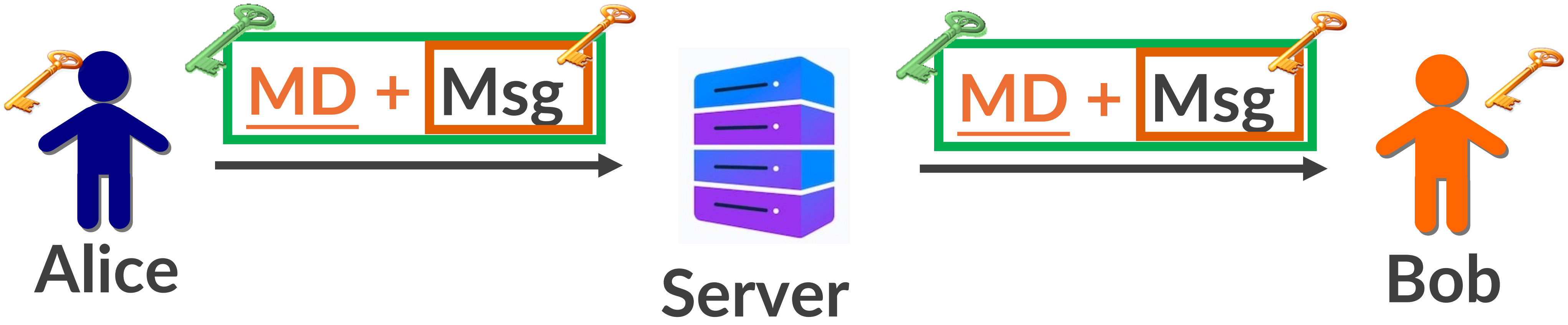
Is Secure Messaging All About E2EE??



No!! In Practice Metadata are Also Used

Metadata: Identity and social relationships of (group) users

(Not considered: login time, geographic locations…



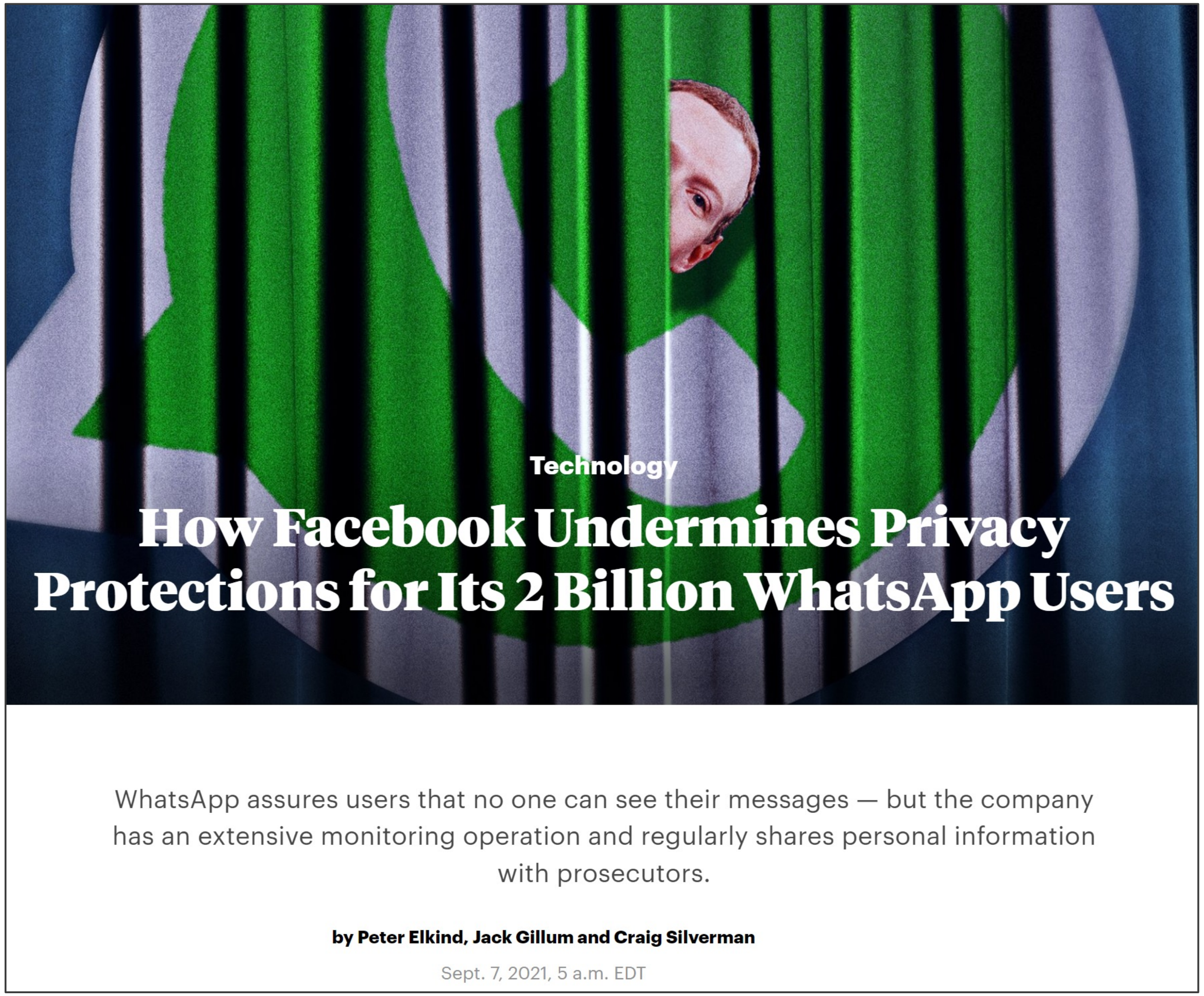
MD := "*Alice to Bob*"

⇒ Server needs MD to relay contents.
But it learns Alice is communicating with Bob ☹

Metadata Reveal More than You Think

“““ If you have enough metadata, you don’t really need content”””

by former NSA general counsel Stewart Baker



Technology

How Facebook Undermines Privacy Protections for Its 2 Billion WhatsApp Users

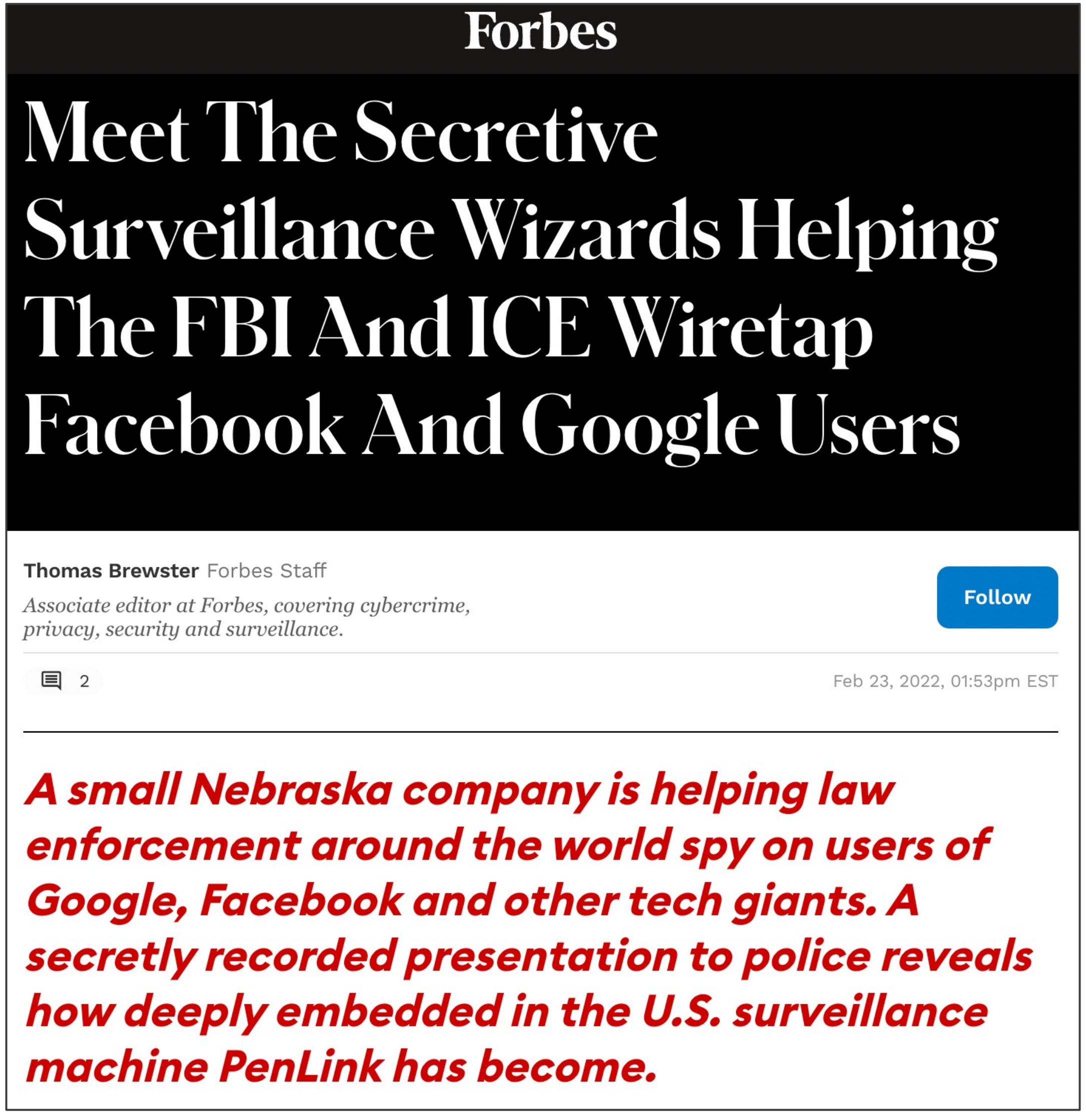
WhatsApp assures users that no one can see their messages — but the company has an extensive monitoring operation and regularly shares personal information with prosecutors.

by Peter Elkind, Jack Gillum and Craig Silverman

Sept. 7, 2021, 5 a.m. EDT

https://www.propublica.org/article/how-facebook-undermines-privacy-protections-for-its-2-billion-whatsapp-users

Metadata Collection is Systematic



Forbes

Meet The Secretive Surveillance Wizards Helping The FBI And ICE Wiretap Facebook And Google Users

Thomas Brewster Forbes Staff

Associate editor at Forbes, covering cybercrime, privacy, security and surveillance.

2

Feb 23, 2022, 01:53pm EST

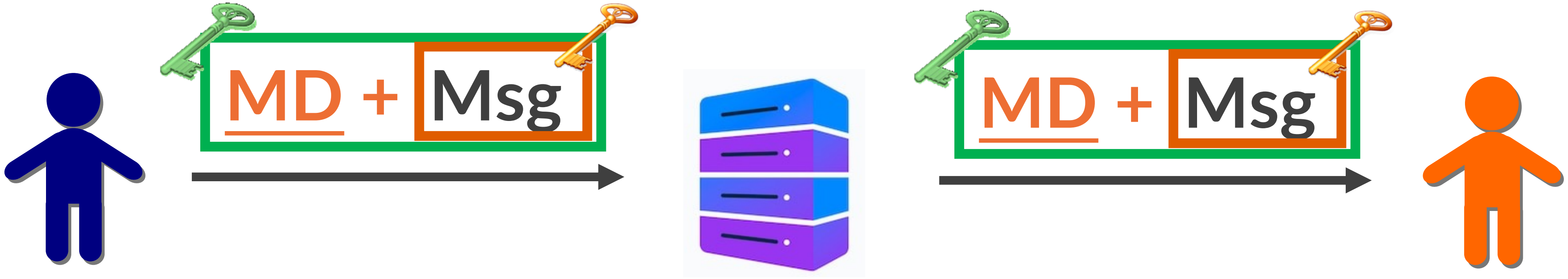
A small Nebraska company is helping law enforcement around the world spy on users of Google, Facebook and other tech giants. A secretly recorded presentation to police reveals how deeply embedded in the U.S. surveillance machine PenLink has become.

> *“““ **Metadata**, however, showing how a WhatsApp account was used and which numbers were contacting one another and when, **can be tracked with a surveillance technology** known as a pen-register. PenLink provides that tool as a service”””*

https://www.forbes.com/sites/thomasbrewster/2022/02/23/meet-the-secretive-surveillance-wiza helping-the-fbi-and-ice-wiretap-facebook-and-google-users/?sh=1d518cff3f0f

Categorizing Sensitive Data into 3 Layers

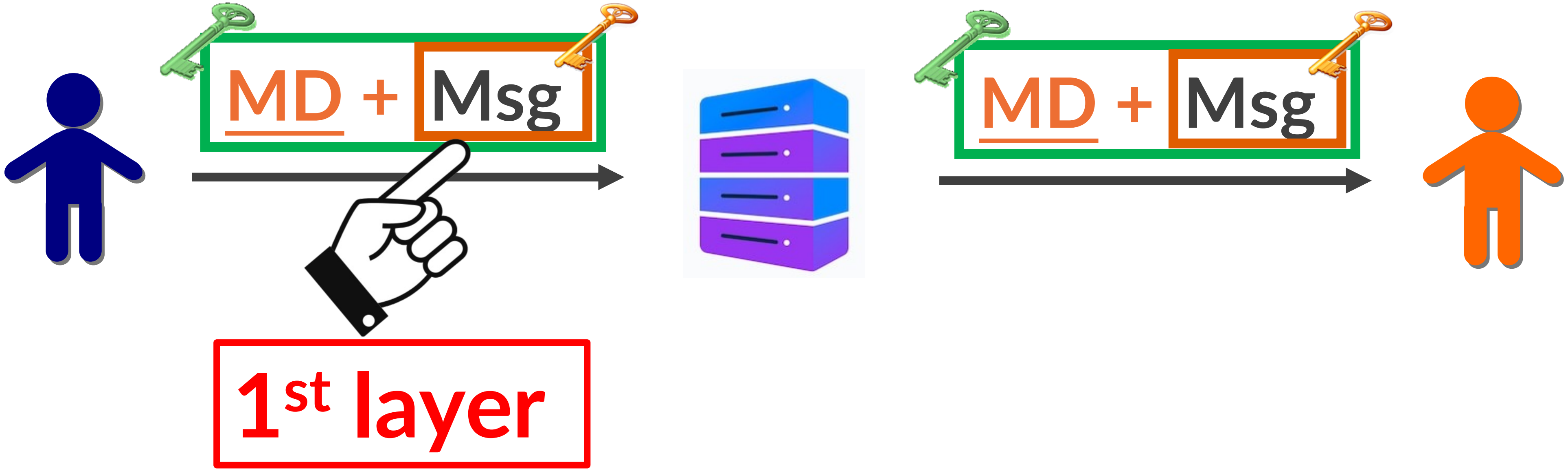
$$\textbf{Sensitive data we want to hide} := \left\{\begin{array}{l}\text{1st layer:}\quad\text{messages}\\ \left.\begin{array}{ll}\text{2nd layer:} & \text{static, explicit metadata}\\ \text{3rd layer:} & \text{dynamic, implicit metadata}\end{array}\right\} =: \text{“metadata”}\end{array}\right.$$



Categorizing Sensitive Data into 3 Layers

Sensitive data we want to hide

$$:= \left\{\begin{array}{ll} \text{1st layer:} & \text{messages} \\ \left.\begin{array}{ll} \text{2nd layer:} & \text{static, explicit metadata} \\ \text{3rd layer:} & \text{dynamic, implicit metadata} \end{array}\right\} & =: \text{“metadata”} \end{array}\right.$$

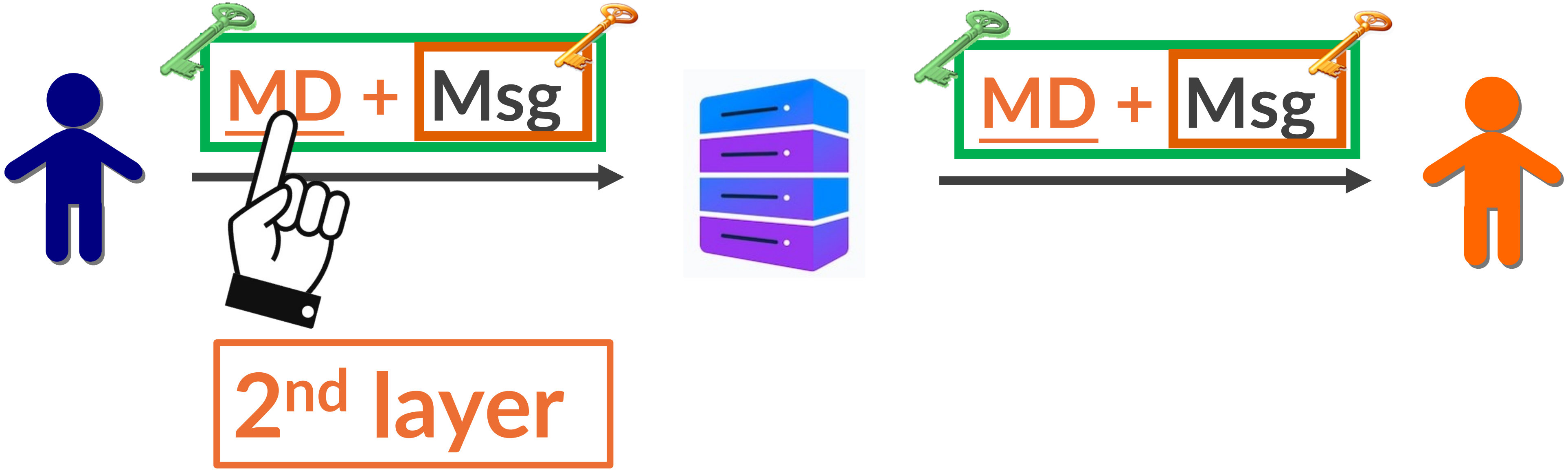


1st layer

Categorizing Sensitive Data into 3 Layers

Sensitive data we want to hide :=

$$\left\{\begin{array}{l} \begin{array}{ll} \text{1st layer:} & \text{messages} \end{array} \\ \left.\begin{array}{ll} \text{2nd layer:} & \text{static, explicit metadata} \\ \text{3rd layer:} & \text{dynamic, implicit metadata} \end{array}\right\} =: \text{“metadata”} \end{array}\right.$$

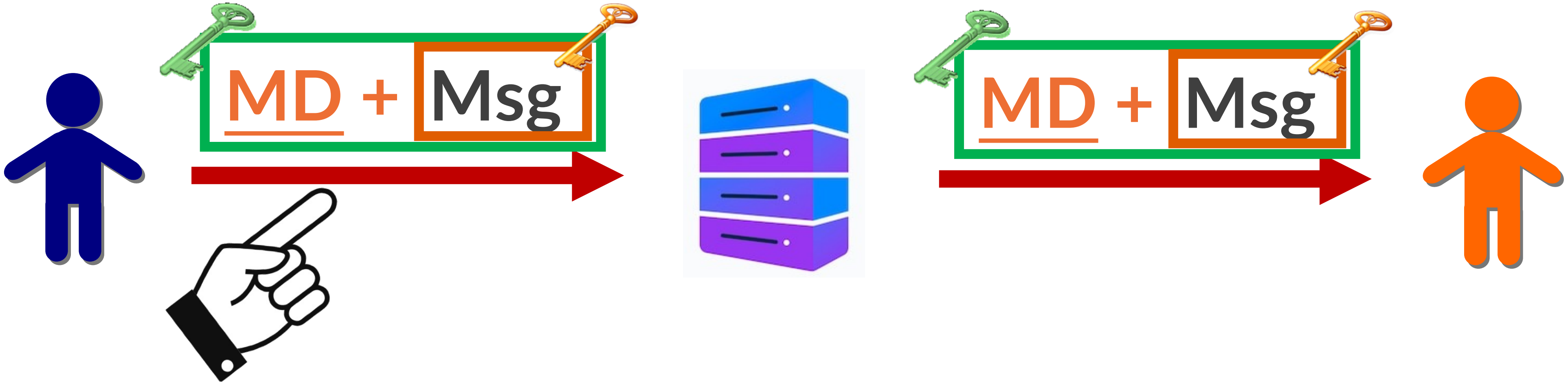


Categorizing Sensitive Data into 3 Layers

Sensitive data we want to hide :=

$$\left\{\begin{array}{ll}\text{1st layer:} & \text{messages} \\ \text{2nd layer:} & \text{static, explicit metadata} \\ \text{3rd layer:} & \text{dynamic, implicit metadata}\end{array}\right.$$

$$\left.\begin{array}{l}\text{static, explicit metadata} \\ \text{dynamic, implicit metadata}\end{array}\right\} =: \text{“metadata”}$$



3rd layer

User-server authentication (e.g., TLS) leaks user identity.
Using anonymous channel allows DoS on users ☹

Current State of Secure Messaging

□ Signal Protocol

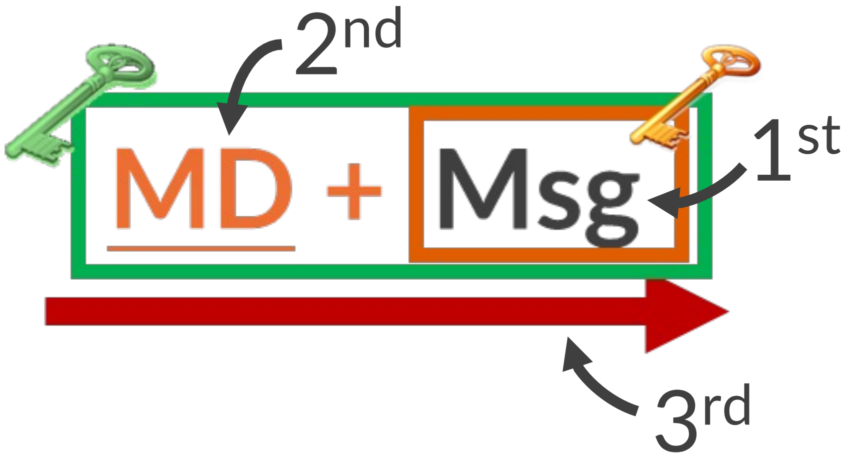
- ✓ Gold standard for two-user messaging.
- ✓ Protocol used within Signal, WhatsApp, Skype…
- ✓ **In the group setting**, sending a message scales with O(N), #users = N ☹

Current State of Secure Messaging

□ Messaging Layer Security (MLS) Protocol

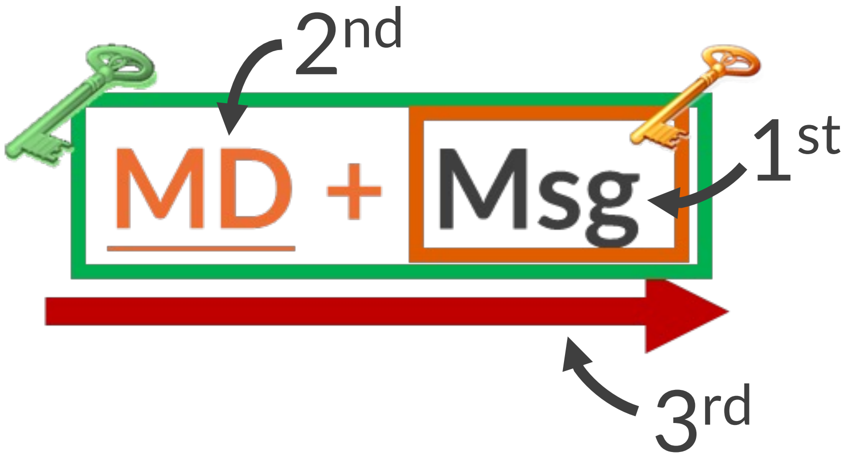
- ✓ Scalable solution to secure messaging.
- ✓ Academic and industry driven IETF is in final stage of standardization.
- ✓ Draft versions of MLS are already used in Cisco’s Webex and RingCentral.
- ✓ Secure as Signal protocol “in spirit”.

Taking a Closer Look



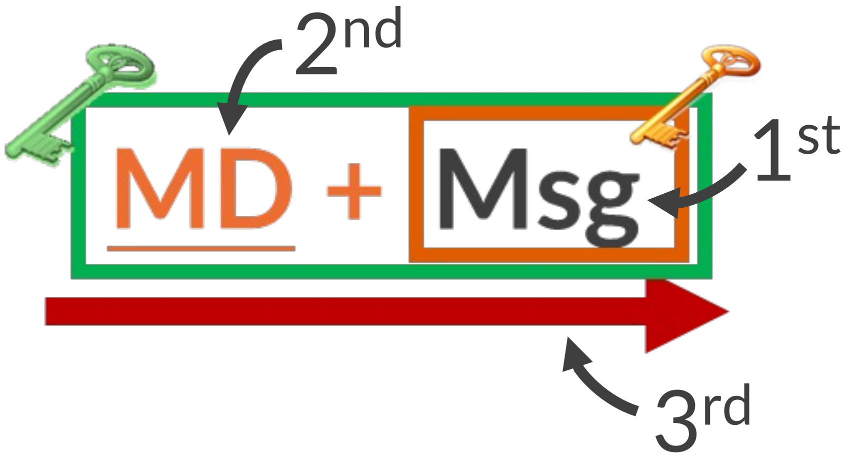
Layer	1st	1st & 2nd	1st,2nd & 3rd
Signal			
Security Proof			
MLS			
Security Proof			

Taking a Closer Look



Layer	1st	1st & 2nd	1st,2nd & 3rd
Signal	Vanilla Signal	2 user: **Sealed Sender** Group: **Private Group**	
Security Proof	2 user: X3DH + DR Group: –	2 user: – Group: [CCS:CPZ20]	
MLS			
Security Proof			

Taking a Closer Look



Layer	1st	1st & 2nd	1st,2nd & 3rd
Signal	Vanilla Signal	2 user: **Sealed Sender** Group: **Private Group**	
Security Proof	2 user: X3DH + DR Group: –	2 user: – Group: [CCS:CPZ20]	
MLS	MLSPlaintext	MLSCiphertext	☠
Security Proof	e.g., [C:AJM22]	☠	

Taking a Closer Look

Making MLS as secure as Signal

*In fact, better PCS guarantee than Signal for metadata ☺

Layer	1st	1st & 2nd	1st,2nd & 3rd
Signal	Vanilla Signal	2 user: **Sealed Sender** Group: **Private Group**	
Security Proof	2 user: X3DH + DR Group: –	2 user: – Group: [CCS:CPZ20]	
MLS	MLSPlaintext	MLSCiphertext	✓ Contrib. 2
Security Proof	e.g., [C:AJM22]	✓ Contrib. 1	✓ Contrib. 3

A Taste of Our Construction

Our Metadata-Hiding MLS

- Use **unique continuously evolving group secret key** shared by the group. (Non-existing in Signal)

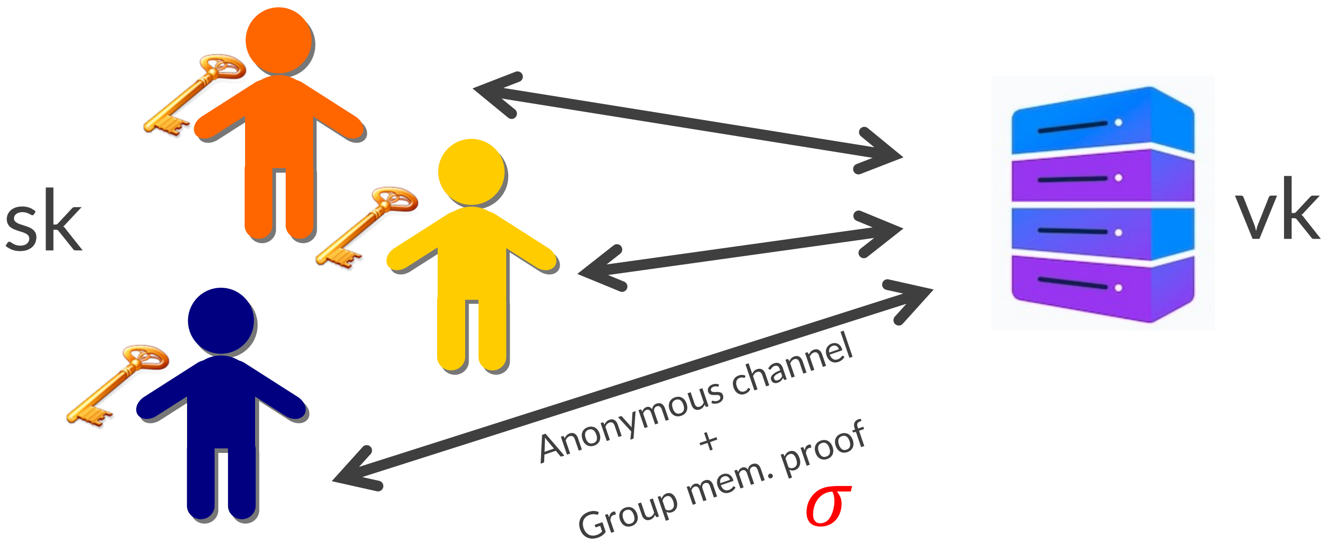
- Bootstraps any MLS-like protocol to be metadata-hiding.

- Generic, simple, efficient, and **post-quantum**.

- **Overhead is only one signature per message ☺**

A Very Simple and Efficient Solution

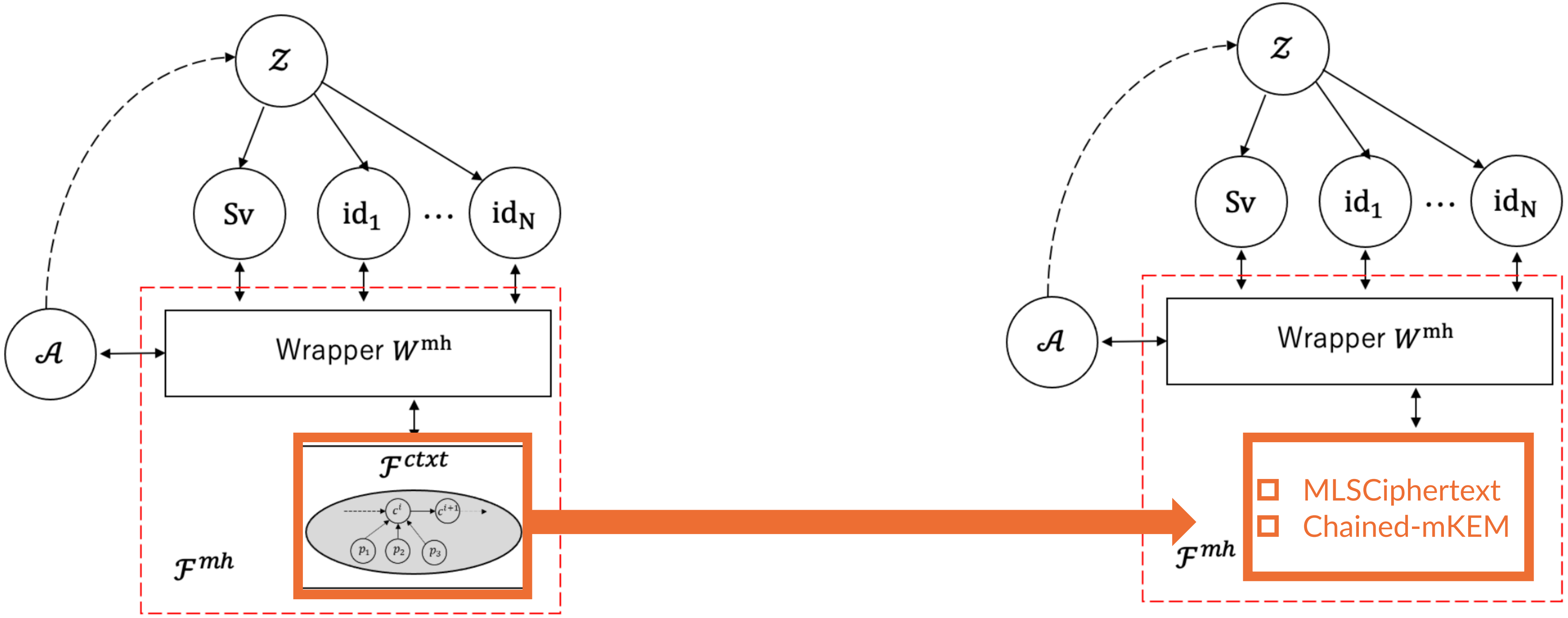
Group secret 🔑 → Derive (vk, sk) for sig. scheme.



- Signing serves as a group membership proof.
- Forward secrecy and post-compromise security of (vk, sk) is inherited from underlying MLS protocol.
- Supports efficient "*selective downloading*" by further deriving secret permutation of group member.

Modular Construction and Security Proof

Our idea works as a wrapper protocol (W^{mh}) for any MLS-like protocol ($\mathcal{F}^{ctxt}$). We formalize this observation in the UC framework ☺



□ MLSCiphertext
□ Chained-mKEM

Ideal MLS-like protocol

Concrete instantiation

Summary

Making MLS-like Protocols (PQ) Metadata Hiding

The Next Step

- We will see more real-world deployment of MLS.
- However, there is a rising tension in regulation of E2EE.
  E.g., EU's Digital Services Act, UK's Online Security Bill
- We also need to work on content moderation.