#### Phoenix: Rebirth of a Cryptographic Password-Hardening Service





**Russell W.F. Lai**<sup>1,2</sup> **Christoph Egger**<sup>1</sup> Dominique Schröder<sup>1</sup> Sherman S.M. Chow<sup>2</sup>

<sup>1</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg
<sup>2</sup>Chinese University of Hong Kong

August 17, 2017

Scheme Design



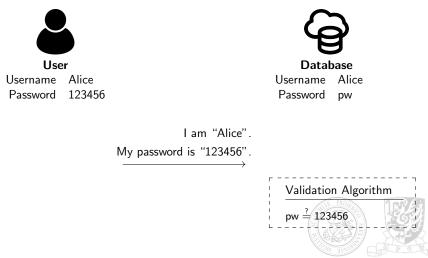
### Password Authentication - before 1976



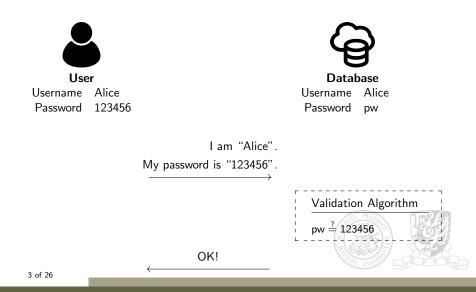
I am "Alice". My password is "123456".



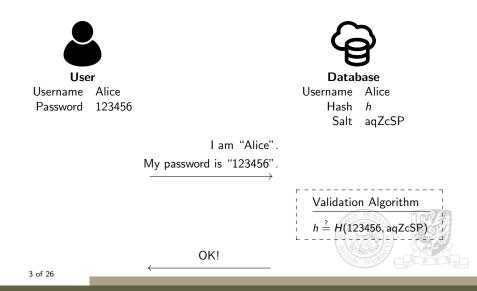
### Password Authentication - before 1976



### Password Authentication - before 1976



### Password Authentication - after 1976



### Problem I - Weak Passwords



Worst Passwords of 2016 by TeamsID

- 4% of users use "123456" as password
- 25% of users use the top 25 worst passwords
- Users are stubborn
  - Choose stronger passwords
- o Use crypto



### Problem II - Stolen Passwords

#### List of data breaches

From Wikipedia, the free encyclopedia

For a broader coverage related to this topic, see Data breach.

This is a list of data breaches, using data compiled from various sources, including press reports, government ne breaches occur continually. Breaches of large organizations where the number of records is still unknown are also

Most breaches occur in North America. It is estimated that the average cost of a data breach will be over \$150 mil breaches <sup>(3)</sup> Vigilante pw@ lists over 2,100 websites which have had their databases breached, containing over 2

Entity •	Year +	Records -
Yahoo	2013	1,000,000,000
Yahoo	2014	500,000,000
Friend Finder Networks	2016	412,214,295
Massive American business hack including 7-Eleven and Nasdaq	2012	160,000,000
Adobe Systems	2013	152,000,000
eBay	2014	145,000,000
Heartland	2009	130,000,000
Rambler.ru	2012	98,167,935
TK / TJ Maxx	2007	94,000,000
AOL	2004	92,000,000

#### Data breaches in 2004-2017 (Wikipedia)



### Problem II - Stolen Passwords

#### List of data breaches

From Wikipedia, the free encyclopedia

For a broader coverage related to this topic, see Data breach

This is a list of data breaches, using data compiled from various sources, including press reports, government ne breaches occur continually. Breaches of large organizations where the number of records is still unknown are also

Most breaches occur in North America. It is estimated that the average cost of a data breach will be over \$150 mil breaches <sup>[3]</sup> Vigilante pw@ lists over 2,100 websites which have had their databases breached, containing over 2

Entity +	Year +	Records -
Yahoo	2013	1,000,000,000
Yahoo	2014	500,000,000
Friend Finder Networks	2016	412,214,295
Massive American business hack including 7-Eleven and Nasdaq	2012	160,000,000
Adobe Systems	2013	152,000,000
eBay	2014	145,000,000
Heartland	2009	130,000,000
Rambler.ru	2012	98,167,935
TK / TJ Maxx	2007	94,000,000
AOL	2004	92,000,000

#### Data breaches in 2004-2017 (Wikipedia)

#### Cost of Data Breach (IBM - 2017 Study)

- Average Cost per Data Breach: \$3.62 million
- Average Cost per Stolen Record \$141



### Problem II - Stolen Passwords

#### List of data breaches

From Wikipedia, the free encyclopedia

For a broader coverage related to this topic, see Data breach

This is a list of data breaches, using data compiled from various sources, including press reports, government ne breaches occur continually. Breaches of large organizations where the number of records is still unknown are also

Most breaches occur in North America. It is estimated that the average cost of a data breach will be over \$150 mil breaches <sup>[3]</sup> Vigilante pw@ lists over 2,100 websites which have had their databases breached, containing over 2

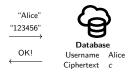
Entity •	Year •	Records -
Yahoo	2013	1,000,000,000
Yahoo	2014	500,000,000
Friend Finder Networks	2016	412,214,295
Massive American business hack including 7-Eleven and Nasdaq	2012	160,000,000
Adobe Systems	2013	152,000,000
eBay	2014	145,000,000
Heartland	2009	130,000,000
Rambler.ru	2012	98,167,935
TK / TJ Maxx	2007	94,000,000
AOL	2004	92,000,000

#### Data breaches in 2004–2017 (Wikipedia)

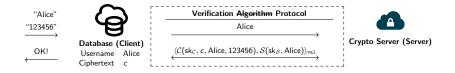
#### Cost of Data Breach (IBM - 2017 Study)

- Average Cost per Data Breach: \$3.62 million
- Average Cost per Stolen Record \$141

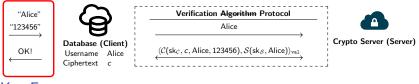
Service providers have incentives to change!







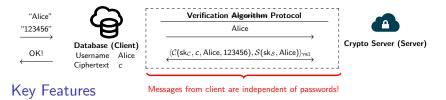




#### Key Features

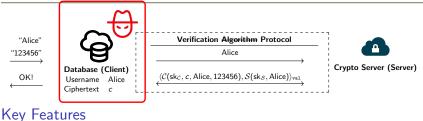
• Seamless to end user (Alice)





- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password

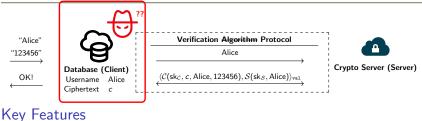




- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

   eliminate offline attacks

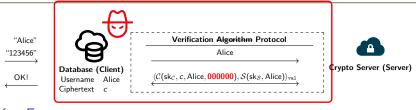




- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

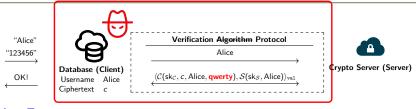
   eliminate offline attacks





- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

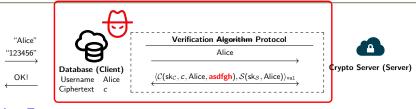
   eliminate offline attacks
- Rate-Limiting (per Username): (Compromised) client cannot submit too many requests
  - mitigate online attacks



- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

   eliminate offline attacks
- Rate-Limiting (per Username): (Compromised) client cannot submit too many requests
  - mitigate online attacks

# Password Hardening Services (Facebook, Pythia)



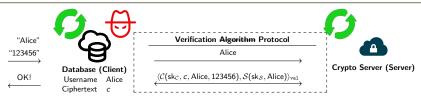
- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

   eliminate offline attacks
- Rate-Limiting (per Username): (Compromised) client cannot submit too many requests
  - mitigate online attacks



- Seamless to end user (Alice)
- Obliviousness: (Malicious) server does not learn password
- Hiding: (Compromised) client cannot verify password by itself

   eliminate offline attacks
- Rate-Limiting (per Username): (Compromised) client cannot submit too many requests
  - mitigate online attacks

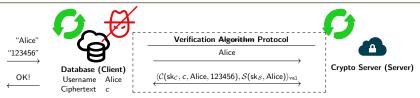


### Key Features

#### Key-Rotation

- Update both keys if either party is compromised
- Bring the entire system to a fresh state
- $\circ~$  Update ciphertexts without knowing passwords (seamless to end user)





### Key Features

#### Key-Rotation

- Update both keys if either party is compromised
- Bring the entire system to a fresh state
- $\circ~$  Update ciphertexts without knowing passwords (seamless to end user)



Russell W. F. Lai

## The Crypto Server



- Key generation independent of client
  - $\circ~$  Can be set up by any third party company / organization
  - One server can serve multiple clients



Russell W. F. Lai

# The Crypto Server



- Key generation independent of client
  - $\circ~$  Can be set up by any third party company / organization
  - One server can serve multiple clients
- Only stores:
  - $\circ~$  One secret key per client
  - One counter per end user for rate-limiting (deleted after the current time interval)



Russell W. F. Lai

# The Crypto Server



- Key generation independent of client
  - $\circ~$  Can be set up by any third party company / organization
  - One server can serve multiple clients
- Only stores:
  - $\circ~$  One secret key per client
  - One counter per end user for rate-limiting (deleted after the current time interval)
- Can be split into multiple servers (future work)



### False Friends (Similar but different notions)

#### Common Feature

• To distribute the task of verifying passwords to multiple servers



### False Friends (Similar but different notions)

#### Common Feature

• To distribute the task of verifying passwords to multiple servers

Distributed Password Verification (CCS'15) Camenisch, Lehmann, and Neven

- · Joint key generation between client and server
- Rate-limiting at client only



### False Friends (Similar but different notions)

#### Common Feature

• To distribute the task of verifying passwords to multiple servers

Distributed Password Verification (CCS'15) Camenisch, Lehmann, and Neven

- · Joint key generation between client and server
- Rate-limiting at client only

Password-Authenticated Key Exchange (PAKE) / Password-Protected Secret Sharing (PPSS)

- Crypto servers need to store a secret share per end user
- No / inefficient key rotations

### Literature on Password Hardening (PH)

Partially-Oblivious Pseudorandom Functions (PO-PRF) (USENIX'15) Everspaugh, Chatterjee, Scott, Juels, and Ristenpart

- Formalized PO-PRF
- Construction (PYTHIA) based on pairing



### Literature on Password Hardening (PH)

Partially-Oblivious Pseudorandom Functions (PO-PRF) (USENIX'15) Everspaugh, Chatterjee, Scott, Juels, and Ristenpart

- Formalized PO-PRF
- Construction (PYTHIA) based on pairing

Partially-Oblivious Commitments (PO-Com) (CCS'16) Schneider, Fleischhacker, Schröder, and Backes

- Formalized PO-Com
  - Security definitions too weak for PH (not covering online attacks)
- Construction without pairing
  - $\circ~2\times$  faster than  $\mathrm{PytHIA}$  when used for PH



- Finally, formalized PH
  - Revisit and strengthen model of Schneider et al.



- Finally, formalized PH
  - Revisit and strengthen model of Schneider et al.
- Formalized key-rotation



- Finally, formalized PH
  - Revisit and strengthen model of Schneider et al.
- Formalized key-rotation
- Devastating online attacks against scheme of Schneider et al.
  - Attack 1: Enable offline-dictionary attack after one validation request
  - Attack 2: Extract password after one validation request
  - The attacks defeat the purpose of external crypto server
  - $\circ~$  The attacks are outside of their security model



- Finally, formalized PH
  - Revisit and strengthen model of Schneider et al.
- Formalized key-rotation
- Devastating online attacks against scheme of Schneider et al.
  - Attack 1: Enable offline-dictionary attack after one validation request
  - Attack 2: Extract password after one validation request
  - $\circ~$  The attacks defeat the purpose of external crypto server
  - $\circ~$  The attacks are outside of their security model
- Extremely simple construction (still without pairing)
  - $\circ~$  Simple enough for real-world use easy to understand and implement
  - $\circ~$  Proven secure under strengthened security model
  - $\circ$  1.5× faster than scheme of Schneider *et al.*
  - $\circ~3\times$  faster than  $\rm Pythia$



### From Salted Hash to PHOENIX (Intuitive Description)

Database (Client) Username un Hash H(un, pw, n<sub>C</sub>)

Client Nonce n<sub>C</sub>



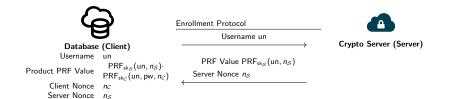
### From Salted Hash to PHOENIX (Intuitive Description)





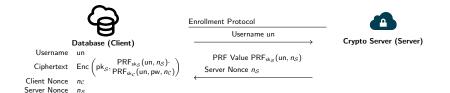
11 of 26

# From Salted Hash to PHOENIX (Intuitive Description)



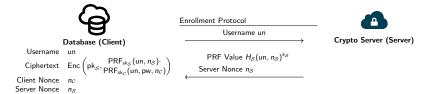


# From Salted Hash to PHOENIX (Intuitive Description)





# From Salted Hash to PHOENIX (Intuitive Description)

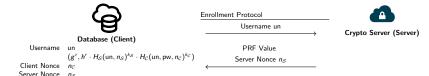


### Homomorphic Encryption Enc(pk, $m \cdot m'$ ) $\approx$ Enc(pk, m) $\cdot$ Enc(pk, m')

### Key-Homomorphic PRF $PRF_{sk \cdot sk'}(m) =$ $PRF_{sk}(m) \cdot PRF_{sk'}(m)$



# From Salted Hash to PHOENIX (Intuitive Description)



e.g., ElGamal (and variants)

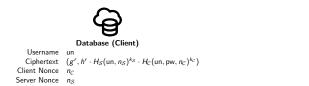
$$\mathsf{sk} = \mathsf{s}, \ \mathsf{pk} = \mathsf{h} = \mathsf{g}^\mathsf{s}$$
  
 $\mathsf{c} = (\mathsf{g}^\mathsf{r}, \mathsf{h}^\mathsf{r} \cdot \mathsf{m})$ 

*e.g.*, Naor-Pinkas-Reingold sk = k  $y = H(m)^k$ 

$$(g^{r+r'}, h^{r+r'} \cdot m \cdot m') = (g^r, h^r \cdot m) \cdot (g^{r'}, h^{r'} \cdot m')$$

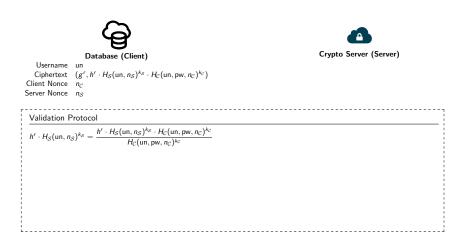
$$H(m)^{k+k'} = H(m)^k \cdot H(m)^{k'}$$

# **PHOENIX** Validation (Intuitive Description)

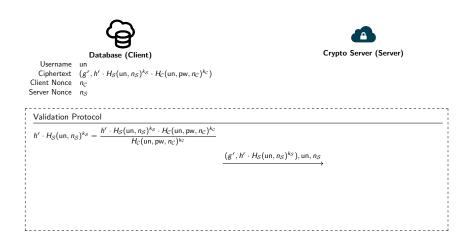




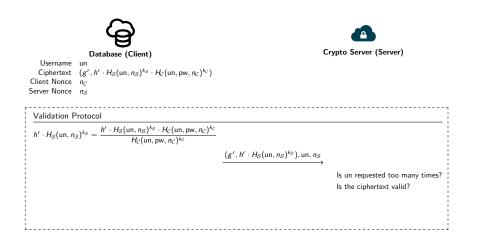
# **PHOENIX** Validation (Intuitive Description)



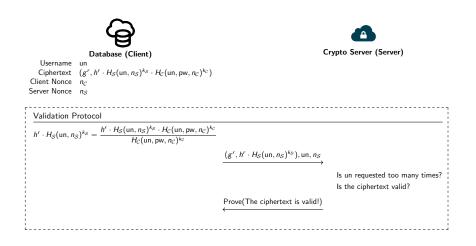
## PHOENIX Validation (Intuitive Description)



### **PHOENIX** Validation (Intuitive Description)



## PHOENIX Validation (Intuitive Description)



# Why it works?

- Obliviousness:
  - $\circ~$  Nothing about the password is sent to the server

$$(g^r, h^r \cdot H_{\mathcal{S}}(\mathrm{un}, n_{\mathcal{S}})^{k_{\mathcal{S}}}), \mathrm{un}, n_{\mathcal{S}})$$



# Why it works?

- Obliviousness:
  - $\circ~$  Nothing about the password is sent to the server

$$(g^r, h^r \cdot H_{\mathcal{S}}(\mathrm{un}, n_{\mathcal{S}})^{k_{\mathcal{S}}}), \mathrm{un}, n_{\mathcal{S}})$$

- Hiding:
  - PRF values of the passwords are encrypted to the server
  - Client cannot decrypt by itself
  - Validity check binds  $(un, n_S)$  with  $H_S(un, n_S)^{k_S}$  in c
  - $\,\circ\,$  Online attacks require guessing pw to remove  $H_{\!\mathcal{C}}(\mathsf{pw},n_{\!\mathcal{C}})^{k_{\!\mathcal{C}}}$  from c



## **PHOENIX** Key-Rotation (Intuitive Description)

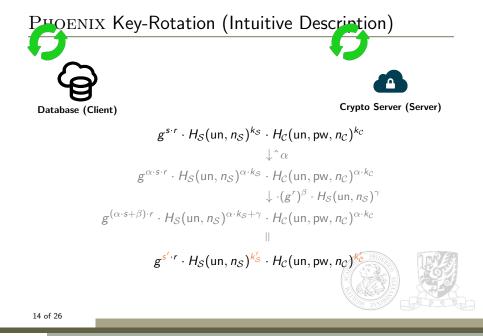




Crypto Server (Server)

 $g^{s \cdot r} \cdot H_{\mathcal{S}}(\mathsf{un}, n_{\mathcal{S}})^{k_{\mathcal{S}}} \cdot H_{\mathcal{C}}(\mathsf{un}, \mathsf{pw}, n_{\mathcal{C}})^{k_{\mathcal{C}}}$ 





#### Evaluation and Deployment



#### In Comparison

Current password hashing recommendations suggest up to one second single-core computing time

#### Context

We have all three (python based) implementations running on Amazon AWS single-core instances

#### Questions

- How long does the user have to wait for password verification
- Do we need many servers to support Password-Hardening
- What are the practical implications

How long must the end user wait for to log in?



#### How long must the end user wait for to log in?

 $\approx$  8 *ms*! (+ round trip time)

	нттр	HTTPS	Frankfurt HTTPS keep-alive	НТТР	HTTPS	Ireland HTTPS keep-alive
RTT (64 bytes)		1.2			23	
$P_{\rm YTHIA} \ enroll/validate$	17.93	25.28	16.01	62.03	113.79	38.56
Schneider et al. enroll	9.80	22.86	8.14	53.72	111.40	30.89
Schneider et al. validate	12.30	25.65	10.73	56.32	115.32	33.49
PHOENIX enroll	5.43	17.93	3.89	50.30	107.25	26.52
PHOENIX validate	9.74	22.78	8.06	53.92	113.02	30.73

Latency in millisecond (ms)



### Evaluation

How many requests can the server entertain in one second?



#### How many requests can the server entertain in one second?

#### Over 370!

	HTTPS	HTTPS
	keep-alive	
parameter	2,607.16	807.50
$\operatorname{Pythia}$ enroll/validate	128.50	125.75
Schneider et al. enroll	380.37	278.51
Schneider <i>et al.</i> validate	221.75	183.92
PHOENIX enroll	1,557.81	697.66
PHOENIX validate	371.34	275.42
		- St Et And

Requests per second

### Practical Deployment

#### Hybrid Scheme

Can we make use of memory-hard functions like Argon2 or scrypt?



#### Hybrid Scheme

Can we make use of memory-hard functions like Argon2 or scrypt?

- Use a memory-hard function instead of a traditional hash function for the PRF
- Even if the attacker has compromised both Client and Server, she has to use the memory-hard function for dictionary attacks

```
Naor-Pinkas-Reingold
sk = k
PRF value H(x)^k
```



#### Hybrid Scheme

Can we make use of memory-hard functions like Argon2 or scrypt?

- Use a memory-hard function instead of a traditional hash function for the PRF
- Even if the attacker has compromised both Client and Server, she has to use the memory-hard function for dictionary attacks

Naor-Pinkas-Reingold sk = kPRF value Sha256(x)<sup>k</sup>



#### Hybrid Scheme

Can we make use of memory-hard functions like Argon2 or scrypt?

- Use a memory-hard function instead of a traditional hash function for the PRF
- Even if the attacker has compromised both Client and Server, she has to use the memory-hard function for dictionary attacks

Naor-Pinkas-Reingold sk = kPRF value  $Argon2(x)^{k}$ 



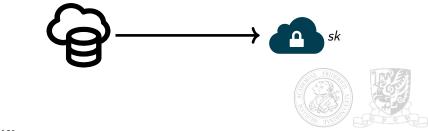
### Practical Deployment

#### Availability



#### Availability

- Server only holds key-pair per Client and Rate-Limiting information
- Several Crypto Servers can host the key-pair for availability **but** keys are then located on several machines



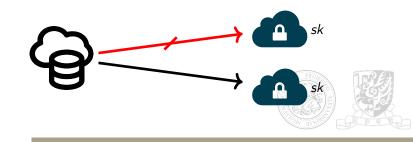
#### Availability

- Server only holds key-pair per Client and Rate-Limiting information
- Several Crypto Servers can host the key-pair for availability **but** keys are then located on several machines



#### Availability

- Server only holds key-pair per Client and Rate-Limiting information
- Several Crypto Servers can host the key-pair for availability **but** keys are then located on several machines



# DDoS and Rate-Limiting

There are two scenarios where Rate-Limiting might be triggered:

Client has been compromised

• It is acceptable when users fail to log in

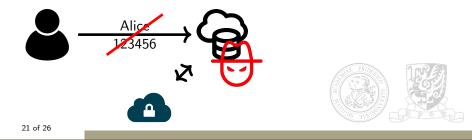


# DDoS and Rate-Limiting

There are two scenarios where Rate-Limiting might be triggered:

Client has been compromised

• It is acceptable when users fail to log in



# DDoS and Rate-Limiting

There are two scenarios where Rate-Limiting might be triggered:

#### External Attacker

- · Honest users should only be slightly inconvenienced
- Crypto Server has little information to distinguish users
- Client is honest and can therefore help

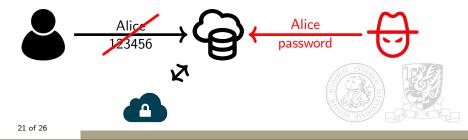


# DDoS and Rate-Limiting

There are two scenarios where Rate-Limiting might be triggered:

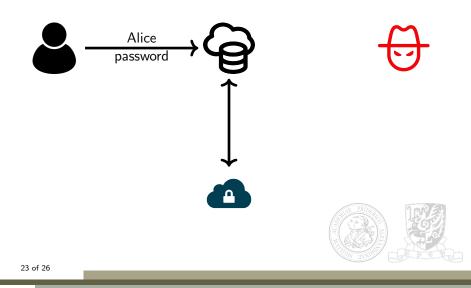
#### External Attacker

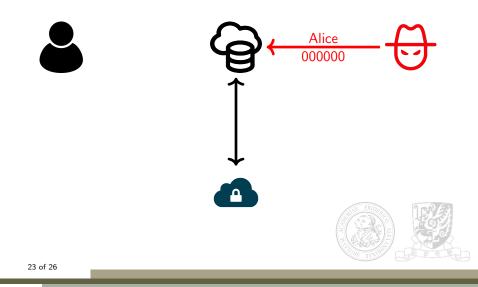
- · Honest users should only be slightly inconvenienced
- Crypto Server has little information to distinguish users
- Client is honest and can therefore help

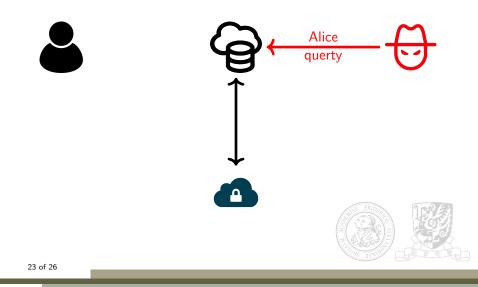


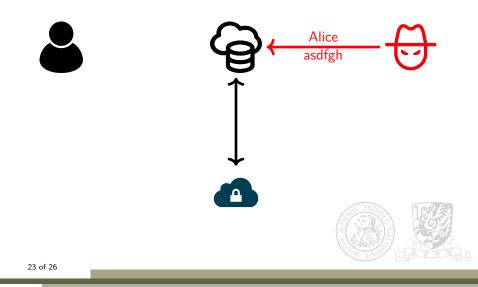
- Warn the Client about upcoming Rate-Limiting: once a soft limit is exceeded there is a limited number of additional tries available the client needs to make sure only the honest user gets to use these
- Client then takes extra measures, for example
  - $\,\circ\,$  Send an E-Mail / SMS /  $\ldots\,$  with an one-time code to the user
  - Add Puzzles to the login screen

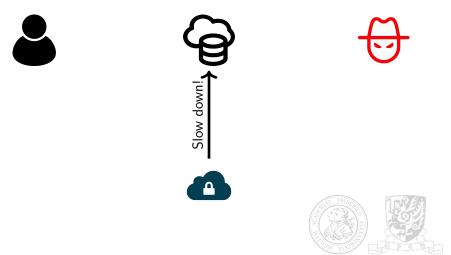












### Rate-Limiting external clients









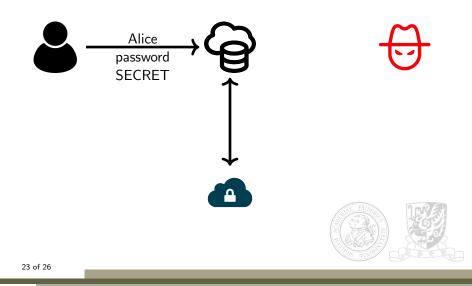












### Upgrade Path

- Both, salted hash and Phoenix need a database field to store data
- Algorithm-ID often already stored alongside the salt and hash
- Upgrade users once they log in

Username	Password	Data
Alice	12:PeADRGbk:gaG4s[]2BwM=	
Bob	12:q79JVDSo:IIRBz[]/9L4=	
Carol	5:3V+ToDHL:FCozKw/gxP/9YZ+Pdr7pcg==	



### Upgrade Path

- Both, salted hash and Phoenix need a database field to store data
- Algorithm-ID often already stored alongside the salt and hash
- Upgrade users once they log in

Username	Password	Data
Alice	12:PeADRGbk:gaG4s[]2BwM=	
Bob	<pre>13:MxTfsL[]phh+8i3DU==</pre>	
Carol	5:3V+ToDHL:FCozKw/gxP/9YZ+Pdr7pcg==	



# Concluding Remark

#### Summary

- Strengthened model of password-hardening
- Formal treatment to key-rotation
- Devastating attack to an existing scheme
- PHOENIX: The most efficient scheme to date



# Concluding Remark

#### Summary

- Strengthened model of password-hardening
- Formal treatment to key-rotation
- Devastating attack to an existing scheme
- PHOENIX: The most efficient scheme to date

#### On Going Projects

- Extended functionality Derive key upon successful validation
- Anonymize end user while retaining rate-limiting
- Deployment by start-up company





Russell W. F. Lai FAU Nuremberg ← Chinese U. Hong Kong



Dominique Schröder FAU Nuremberg



#### Christoph Egger FAU Nuremberg



Sherman S. M. Chow Chinese U. Hong Kong