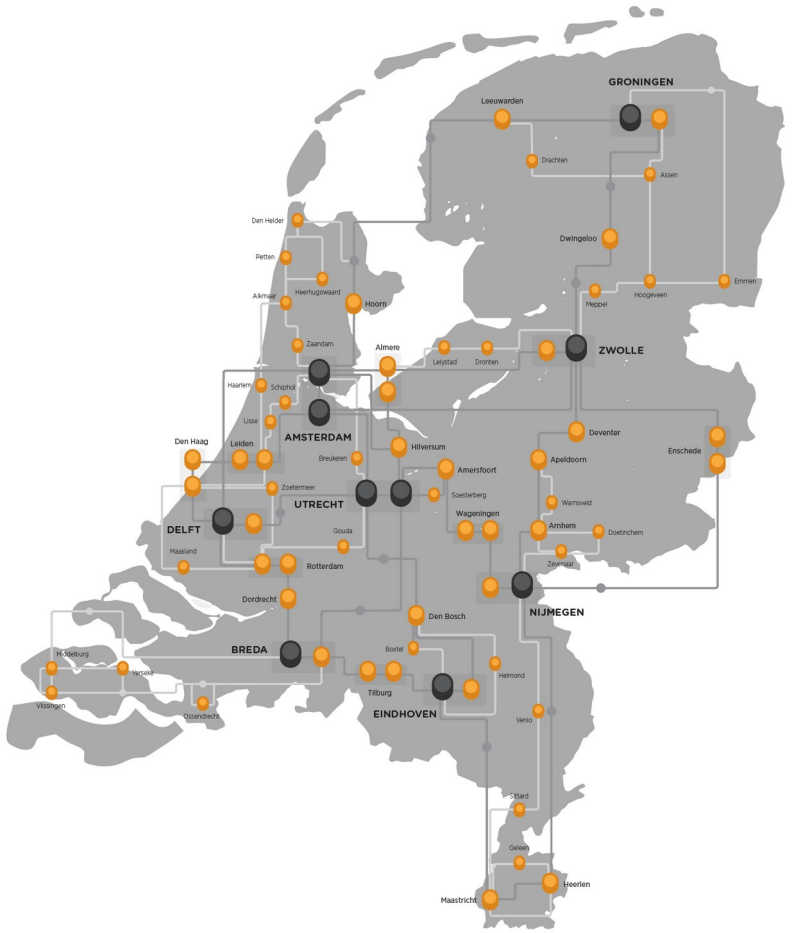
PRIVACY-VRIENDELIJK LOGGEN VOOR INCIDENT-RESPONSE

Joeri de Ruiter

SURF

- De ict-coöperatie van onderwijs en onderzoek
- > 100 leden: wo, hbo, mbo, umc en onderzoek
- Dienstverlening, innovatiewerkplaats, vereniging
- NREN: National Research and Education Network
  - ~12.000 km glasvezel
  - 1,5 miljoen eindgebruikers



Hoe het begon

- Ransomware aanval op een aangesloten instelling
- Lijst met domeinnamen die konden wijzen op een besmetting
- Domeinnamen gedeeld met andere instellingen
- Vraag aan SURF:
  - Wie heeft deze domeinnamen opgevraagd?
  - Geen idee: dit houden we niet bij
- **Hoe kunnen we achterhalen wie een bepaalde domeinnaam heeft opgevraagd?**

Waarom willen we dit?

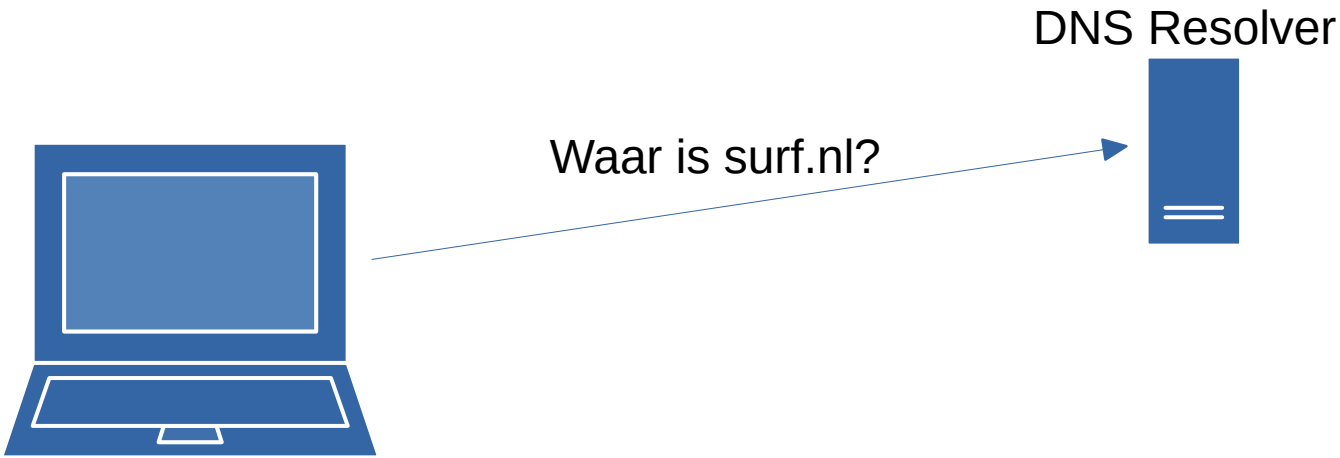
- Detectie van besmette systemen
  - Malware vraagt vaak specifieke domeinnamen op
  - Bijvoorbeeld voor verbinding met command & control servers
- Real-time alerts
  - Bekende kwaadaardige domeinnamen
- Terugzoeken
  - Nieuw gevonden kwaadaardige domeinnamen

DNS intro

- Telefoonboek van het internet
- Vertaling domeinnamen naar IP addressen
- Voorbeeld: surf.nl → 2001:610:508:108:192:87:108:15

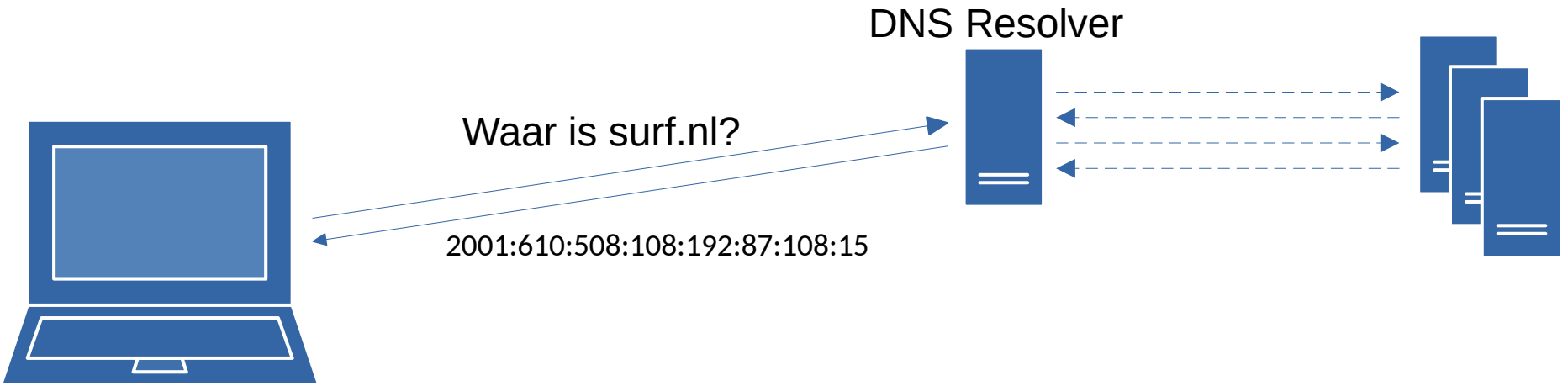
DNS intro

- Telefoonboek van het internet
- Vertaling domeinnamen naar IP addressen
- Voorbeeld: surf.nl → 2001:610:508:108:192:87:108:15



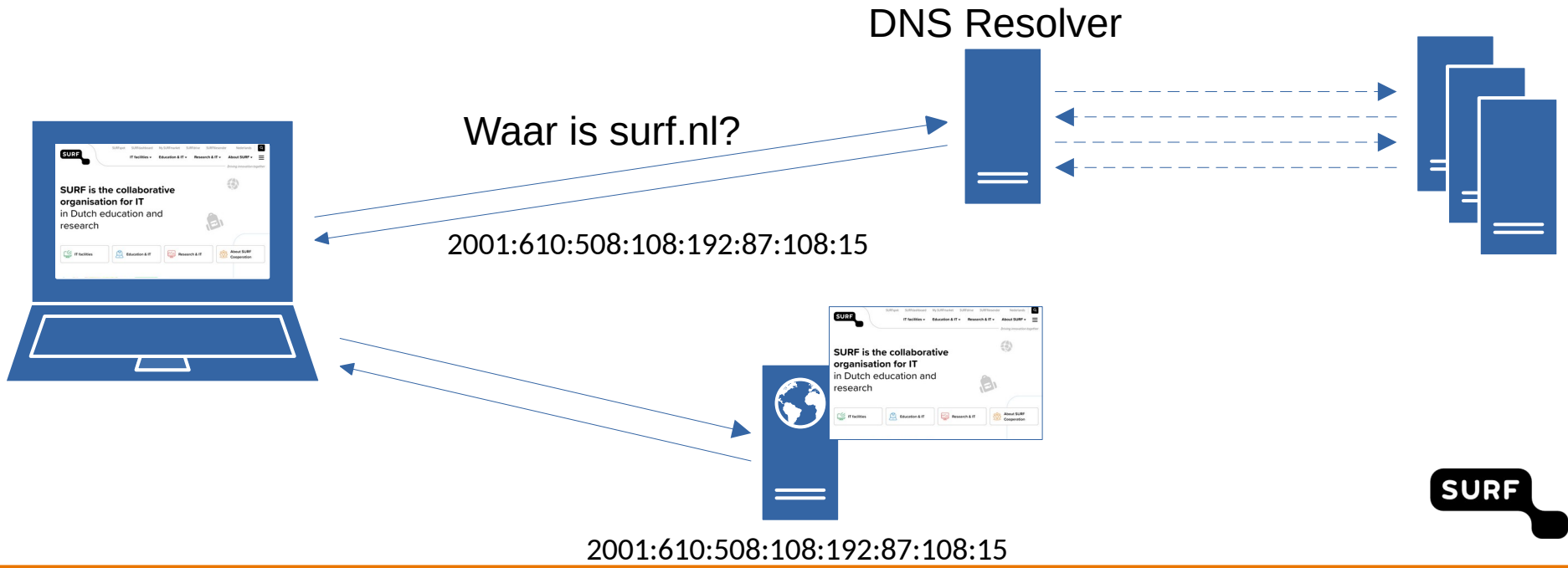
DNS intro

- Telefoonboek van het internet
- Vertaling domeinnamen naar IP addressen
- Voorbeeld: surf.nl → 2001:610:508:108:192:87:108:15



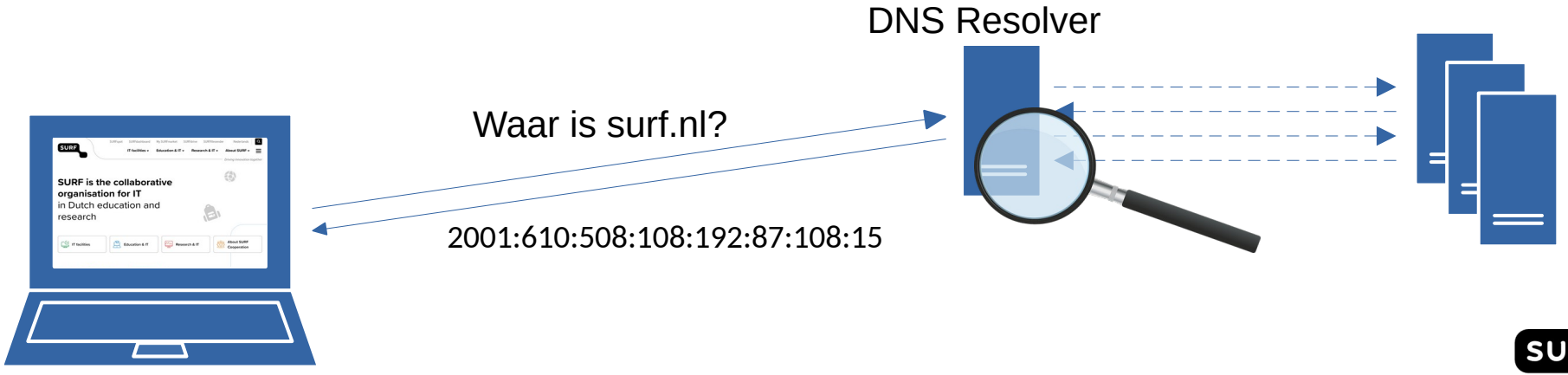
DNS intro

- Telefoonboek van het internet
- Vertaling domeinnamen naar IP addressen
- Voorbeeld: surf.nl → 2001:610:508:108:192:87:108:15



Wie heeft een bepaalde domeinnaam opgevraagd?

- Log alle opvragingen op de DNS resolver
- Erg nuttig om besmette systemen te vinden
  - Kwaadaardige domeinnamen niet altijd meteen bekend
- Echter: zeer privacy-gevoelig
  - De queries bevatten browsegedrag van een gebruiker



Wat willen we weten?

- Welke gebruikers hebben een kwaadaardige domeinnaam opgevraagd?
- Niet: welke domeinnamen heeft een bepaalde gebruiker opgevraagd?

Wat willen we weten?

- Welke gebruikers hebben een kwaadaardige domeinnaam opgevraagd?
- Niet: welke domeinnamen heeft een bepaalde gebruiker opgevraagd?
- Dezelfde data nodig
  - Domeinnaam
  - Informatie over de gebruiker
- Eenvoudige oplossing: zet alles in een database
  - Niet echt privacy-vriendelijk

Wat willen we weten?

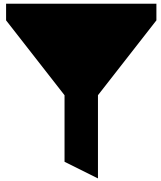
- Welke gebruikers hebben een kwaadaardige domeinnaam opgevraagd?
- Niet: welke domeinnamen heeft een bepaalde gebruiker opgevraagd?
- Dezelfde data nodig
  - Domeinnaam
  - Informatie over de gebruiker
- Eenvoudige oplossing: zet alles in een database
  - Niet echt privacy-vriendelijk
- **Hoe kunnen we DNS opvragingen loggen op een privacy-vriendelijke manier?**
  - **Kunnen we een systeem ontwerpen waarin we wel de eerste, maar niet de tweede vraag kunnen beantwoorden?**

Intermezzo: cryptografische hashfuncties

- Zet tekst om in een willekeurig lijkende code
  - Altijd dezelfde code voor dezelfde tekst
  - Veschillende codes voor verschillende teksten
  - Codes hebben altijd dezelfde lengte
  - Met een gegeven code is het niet eenvoudig te achterhalen wat de oorspronkelijke tekst was

Intermezzo: cryptografische hashfuncties

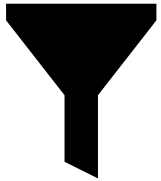
surf.nl



51EF8FD9D7B9996C89DB2F0AD3F91C50

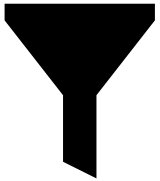
Intermezzo: cryptografische hashfuncties

surf.nl



51EF8FD9D7B9996C89DB2F0AD3F91C50

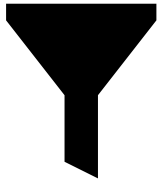
surfdomeinen.nl



CEAC9BF76D801DAB8F5C9889472DF84A

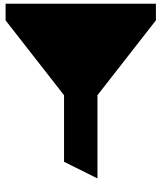
Intermezzo: cryptografische hashfuncties

surf.nl



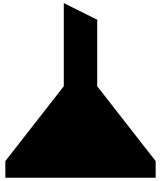
51EF8FD9D7B9996C89DB2F0AD3F91C50

surfdomeinen.nl



CEAC9BF76D801DAB8F5C9889472DF84A

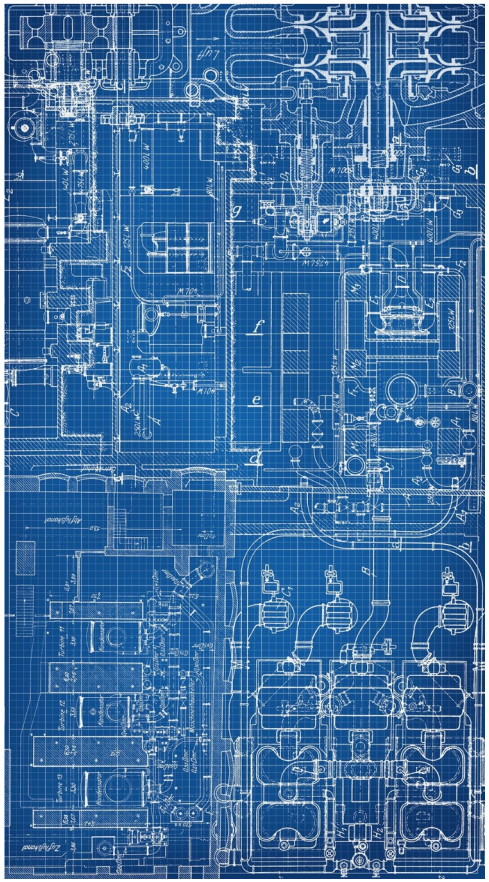
F20404BDD1003CDF110C929F8C238C98



?

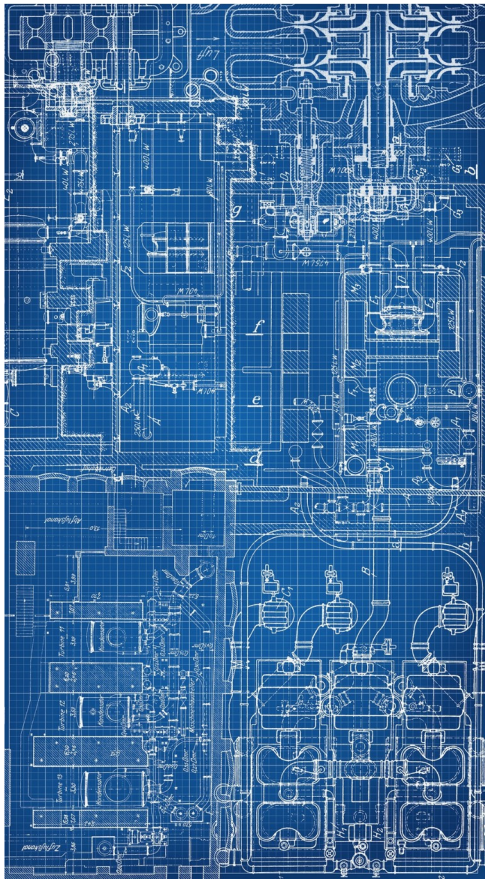
Privacy-vriendelijke opslag

- Sla de hash van de opvraagde domeinnaam op
  - Domeinnaam zoeken mogelijk
  - Extra beveiliging door sleutel bij hash



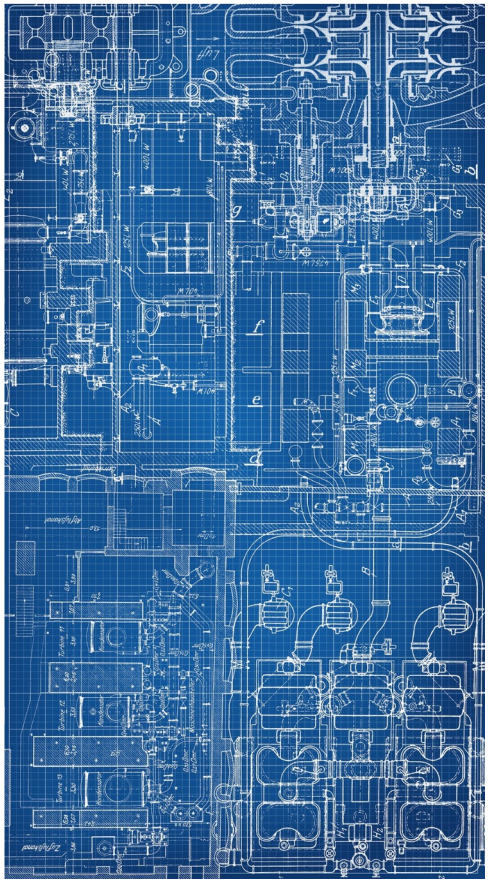
Privacy-vriendelijke opslag

- Sla de hash van de opvraagde domeinnaam op
  - Domeinnaam zoeken mogelijk
  - Extra beveiliging door sleutel bij hash
- Informatie over de gebruiker versleuteld
  - Sleutel gebaseerd op domeinnaam
  - Alleen te ontsleutelen als bijbehorende domeinnaam bekend is



Privacy-vriendelijke opslag

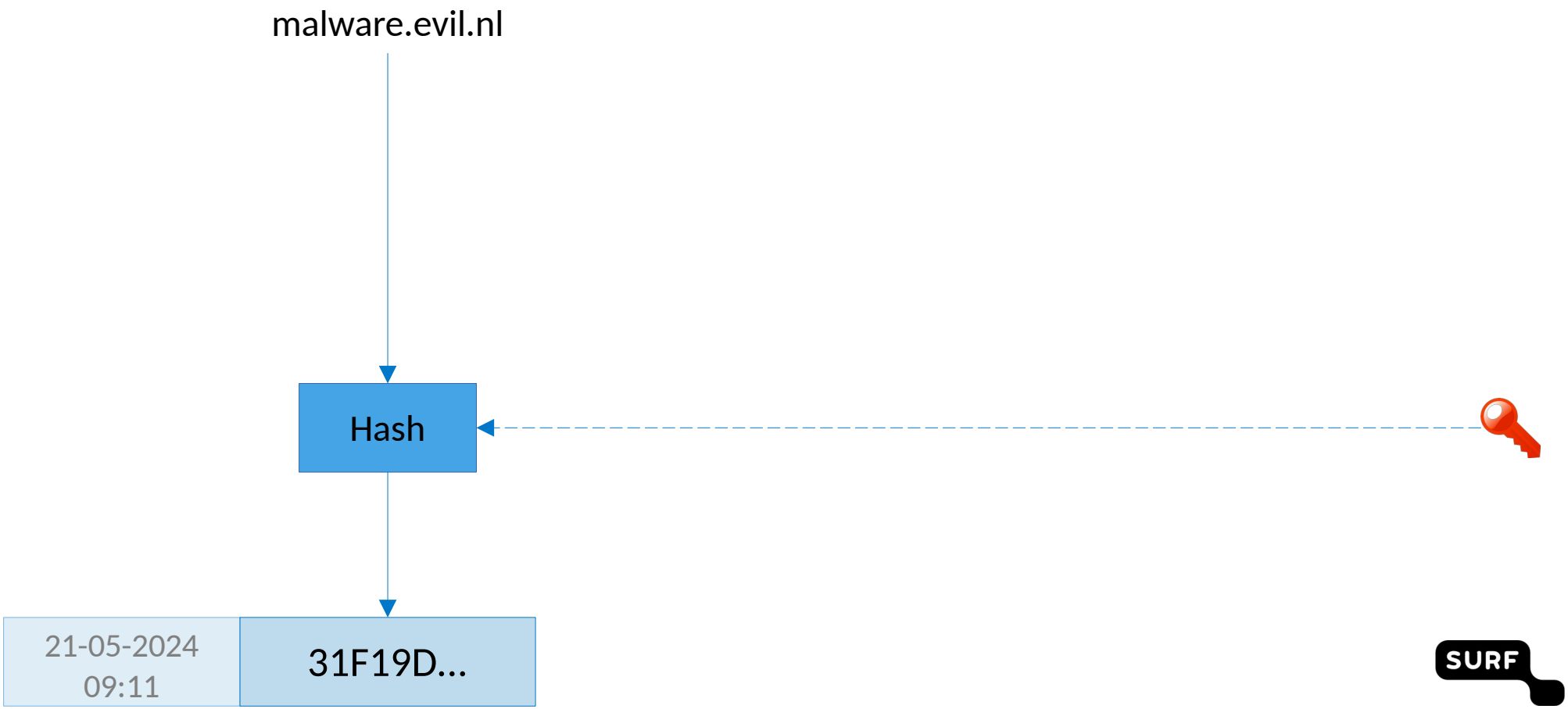
- Sla de hash van de opvraagde domeinnaam op
  - Domeinnaam zoeken mogelijk
  - Extra beveiliging door sleutel bij hash
- Informatie over de gebruiker versleuteld
  - Sleutel gebaseerd op domeinnaam
  - Alleen te ontsleutelen als bijbehorende domeinnaam bekend is
- Domeinnaam zelf niet opgeslagen
  - Alleen afgeleide informatie
  - Niet mogelijk om te zien welke domeinnamen een bepaalde gebruiker heeft opgevraagd



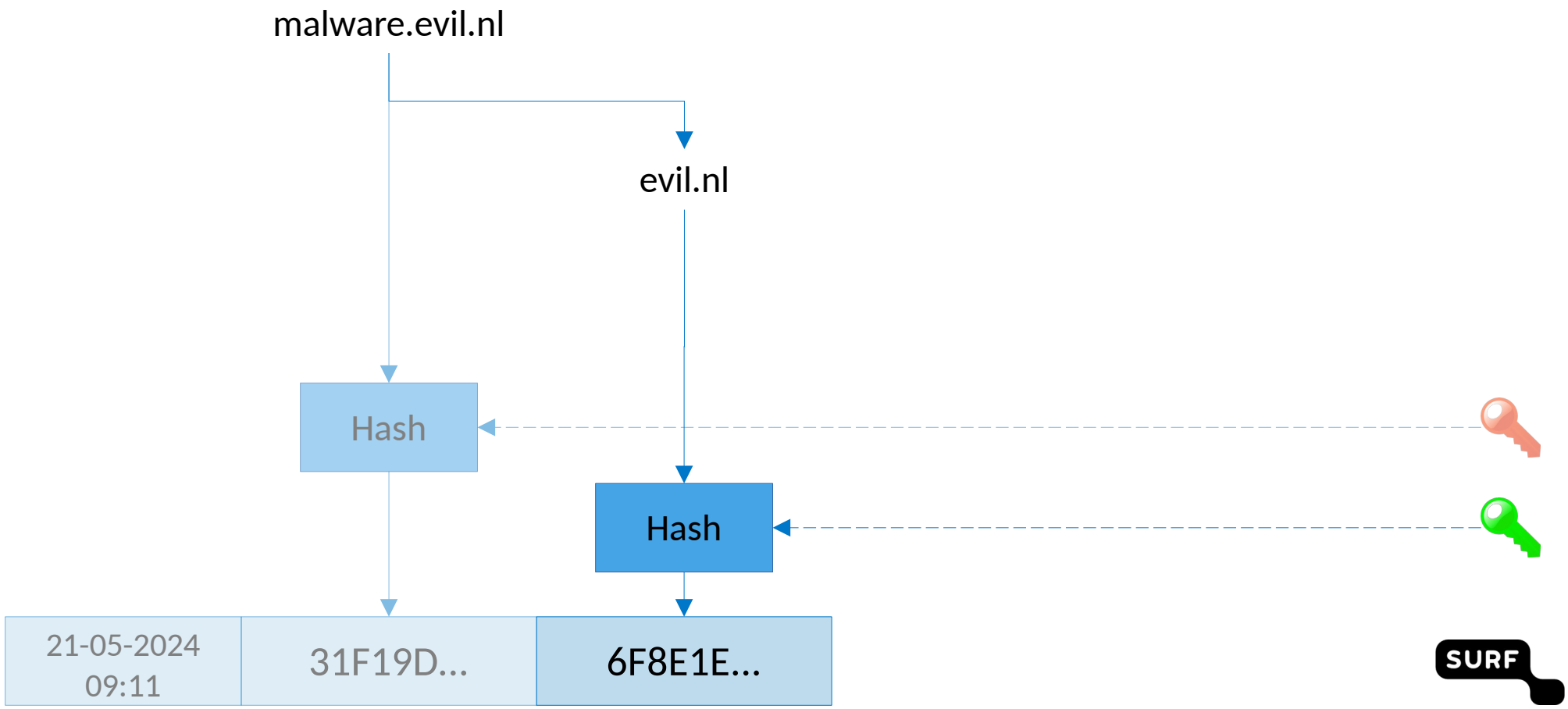
192.0.2.123 vraagt malware.evil.nl op

21-05-2024
09:11

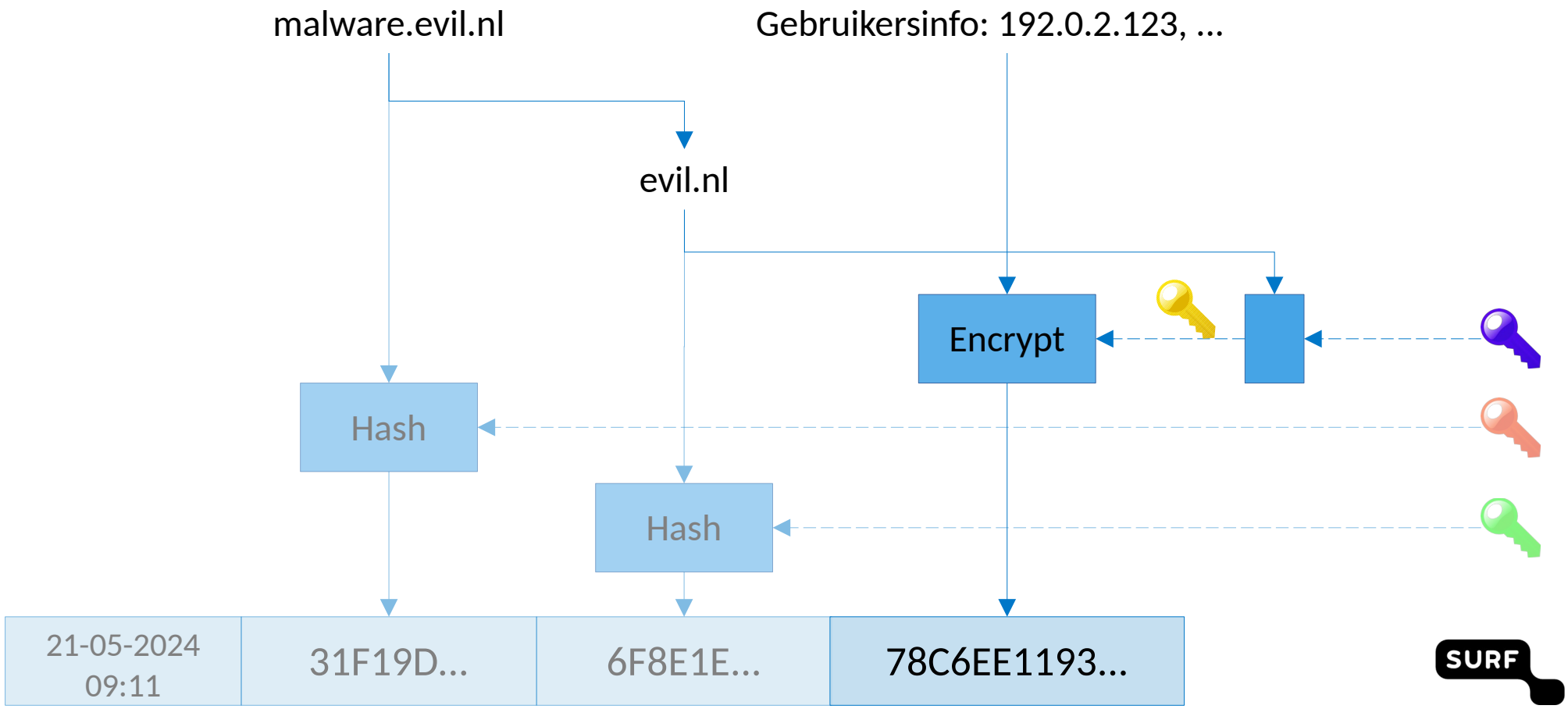
192.0.2.123 vraagt malware.evil.nl op



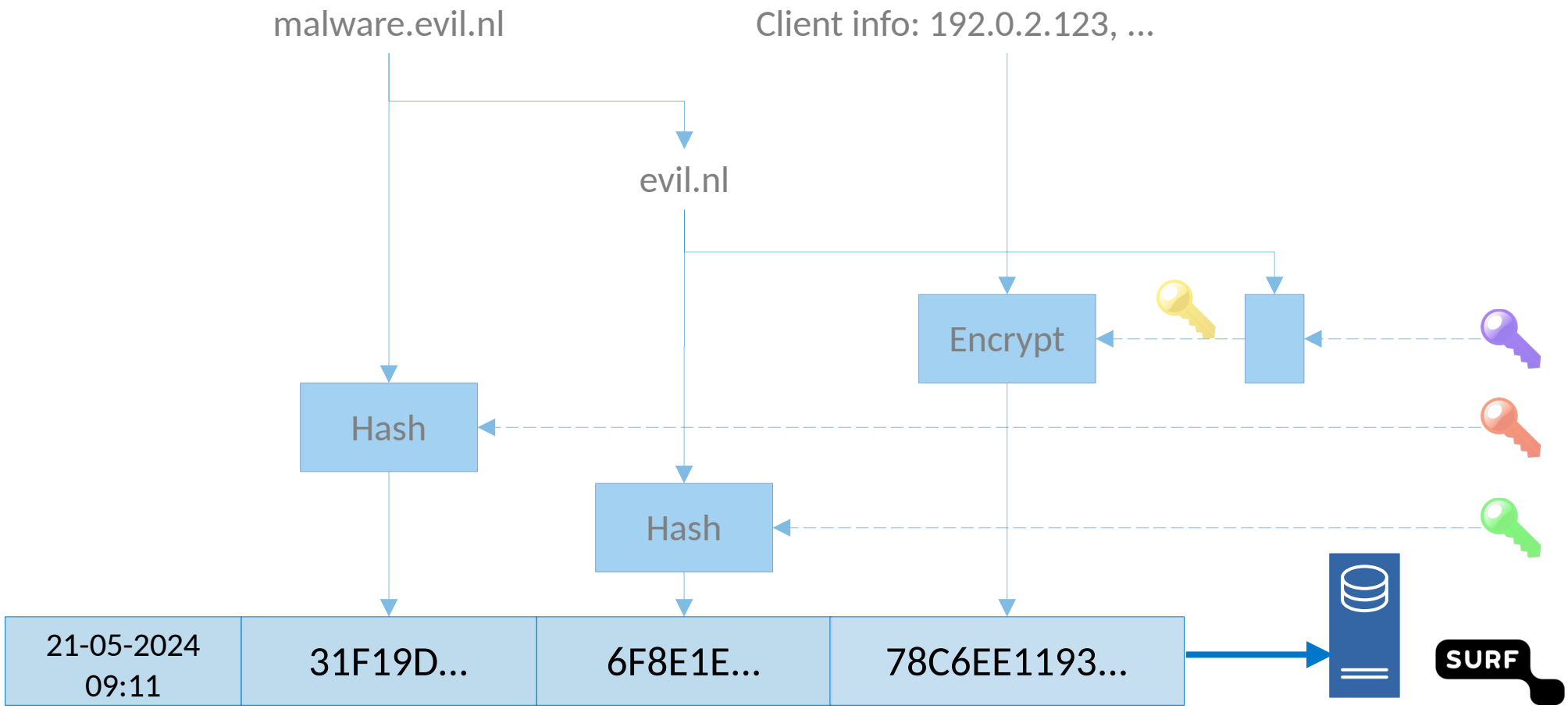
192.0.2.123 vraagt malware.evil.nl op



192.0.2.123 vraagt malware.evil.nl op



192.0.2.123 vraagt malware.evil.nl op

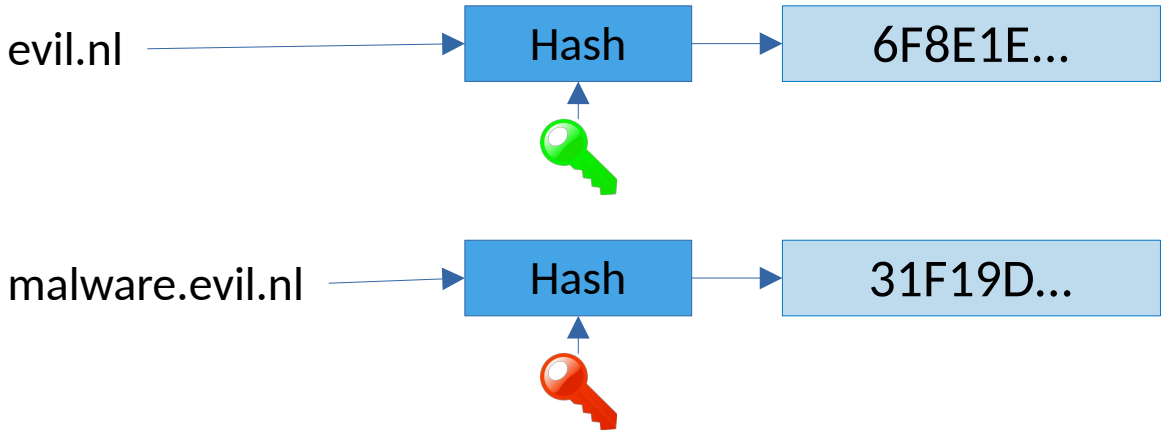


Wie heeft malware.evil.nl opgevraagd?

Tijd	Domeinnaam		Gebruikersinfo
21-05-2024 09:05	A623BE…	E47422…	DF55912A1B…
21-05-2024 09:10	38FE31…	12A8B3…	56ED4CB784…
21-05-2024 09:11	31F19D…	6F8E1E…	78C6EE1193…
21-05-2024 09:14	6789DE…	BF5136…	C45DE823FF…

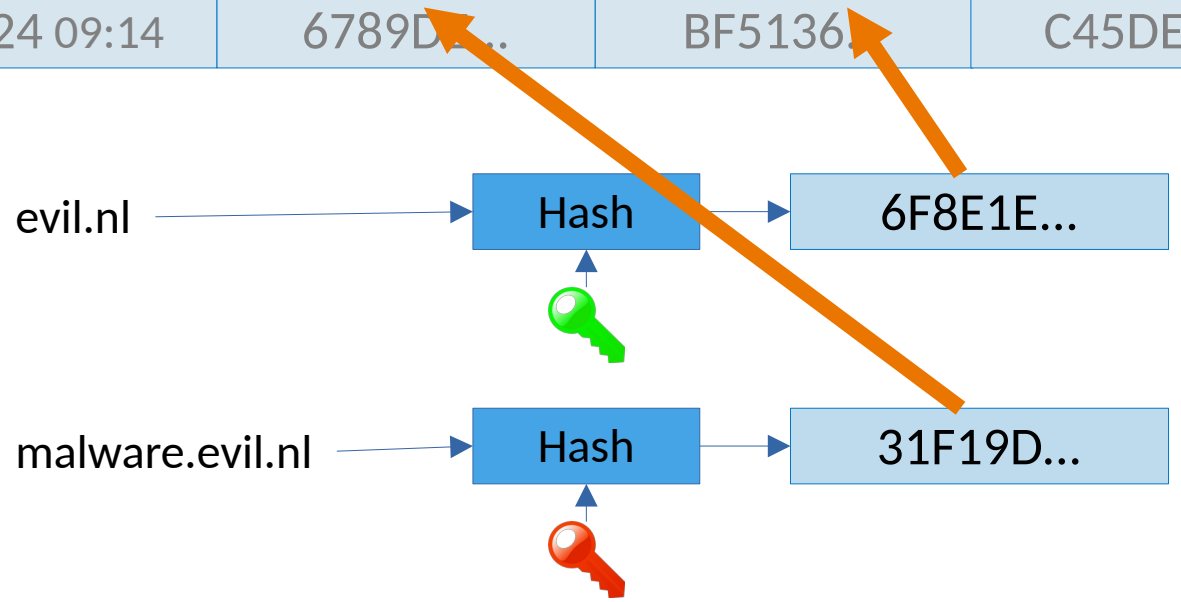
Wie heeft malware.evil.nl opgevraagd?

Tijd	Domeinnaam		Gebruikersinfo
21-05-2024 09:05	A623BE…	E47422…	DF55912A1B…
21-05-2024 09:10	38FE31…	12A8B3…	56ED4CB784…
21-05-2024 09:11	31F19D…	6F8E1E…	78C6EE1193…
21-05-2024 09:14	6789DE…	BF5136…	C45DE823FF…



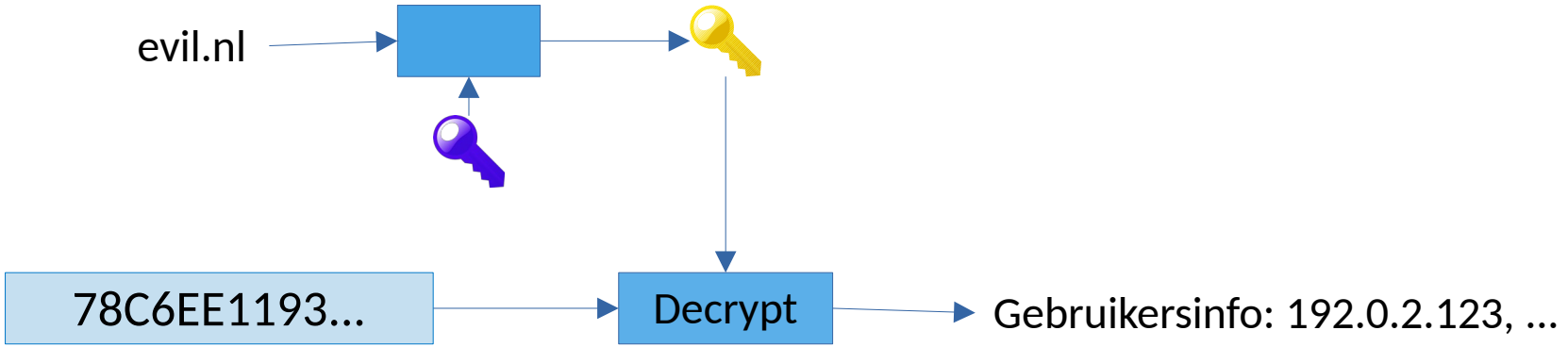
Wie heeft malware.evil.nl opgevraagd?

Tijd	Domeinnaam		Gebruikersinfo
21-05-2024 09:05	A623BE…	E47422…	DF55912A1B…
21-05-2024 09:10	38FE31…	12A8B3…	56ED4CB784…
21-05-2024 09:11	31F19D…	6F8E1E…	78C6EE1193…
21-05-2024 09:14	6789D…	BF5136…	C45DE823FF…



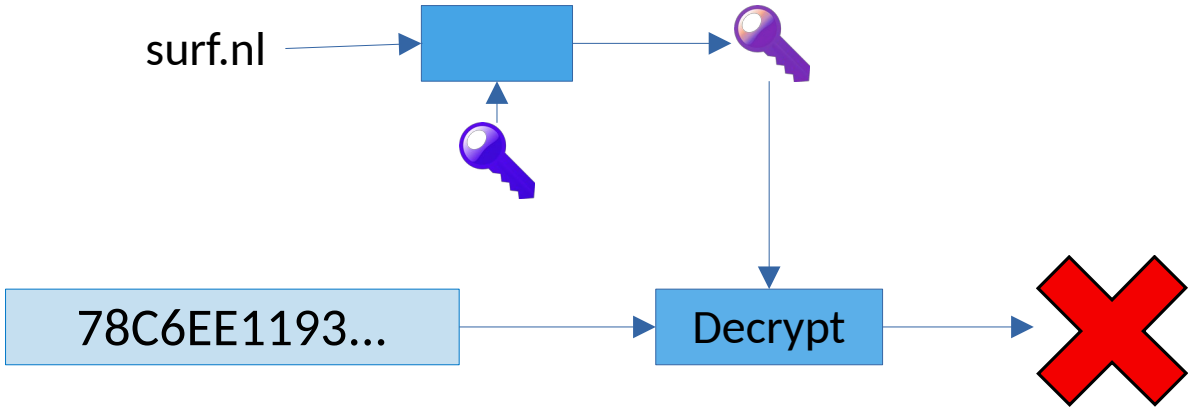
Wie heeft malware.evil.nl opgevraagd?

Tijd	Domeinnaam		Gebruikersinfo
21-05-2024 09:05	A623BE…	E47422…	DF55912A1B…
21-05-2024 09:10	38FE31…	12A8B3…	56ED4CB784…
21-05-2024 09:11	31F19D…	6F8E1E…	78C6EE1193…
21-05-2024 09:14	6789DE…	BF5136…	C45DE823FF…



Wie heeft malware.evil.nl opgevraagd?

Tijd	Domeinnaam		Gebruikersinfo
21-05-2024 09:05	A623BE…	E47422…	DF55912A1B…
21-05-2024 09:10	38FE31…	12A8B3…	56ED4CB784…
21-05-2024 09:11	31F19D…	6F8E1E…	78C6EE1193…
21-05-2024 09:14	6789DE…	BF5136…	C45DE823FF…

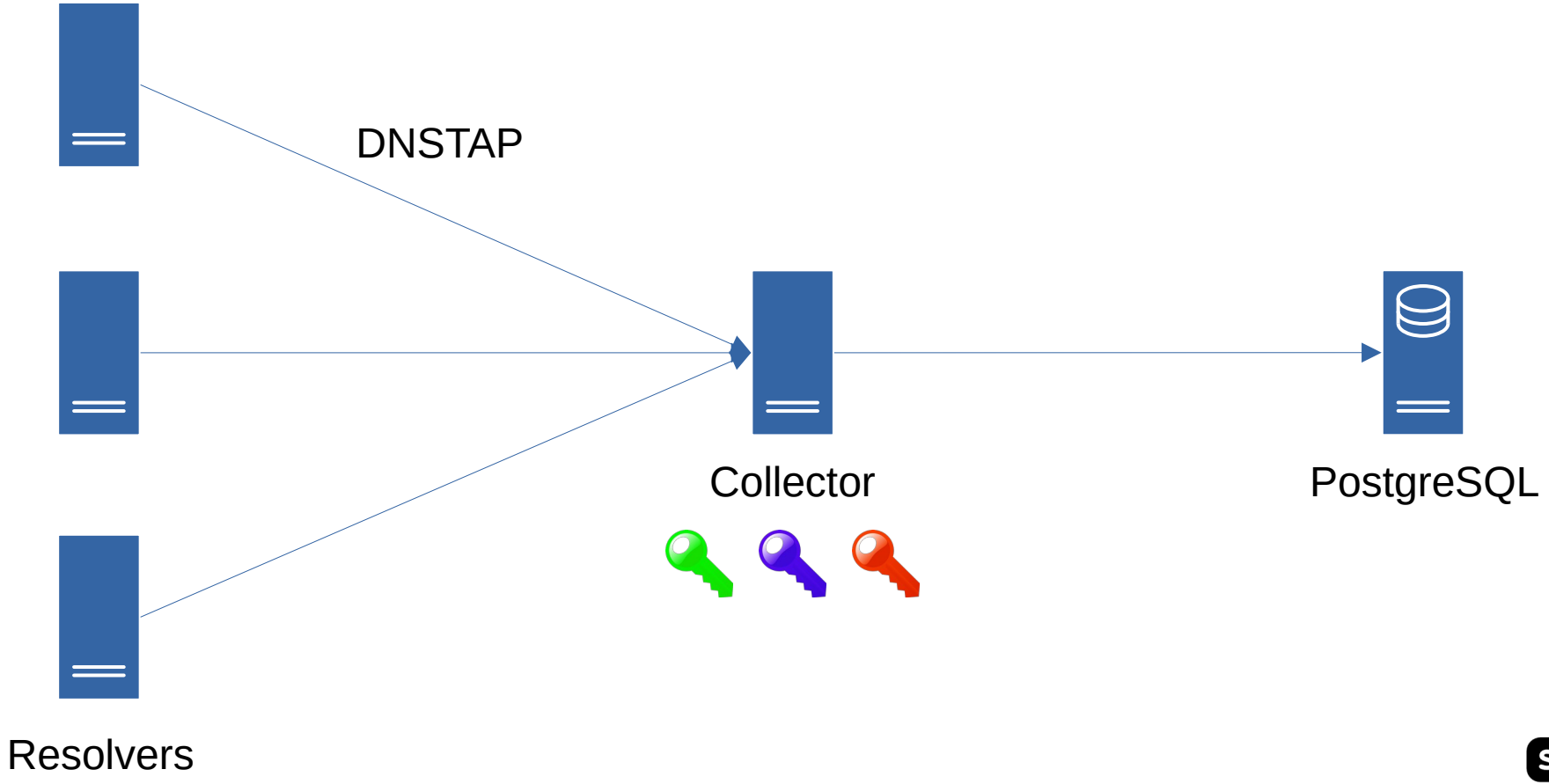


Privacy-vriendelijke opslag

- Wie heeft malware.evil.nl opgevraagd?
  - Sleutels nodig

- Welke domeinnamen heeft 192.0.2.123 opgevraagd?
  - Gebruikersinformatie is versleuteld met behulp van de opgevraagde domeinnaam

Setup



DNSTAP

- DNS queries verzameld met behulp van DNSTAP
- Binair logformaat voor DNS berichten
  - Protocol Buffers
- Ondersteund door de meeste DNS servers
- Non-blocking
  - Minimale impact op de resolver

DNSTAP

```
message Message {
    required Type               type = 1;
    optional SocketFamily       socket_family = 2;
    optional SocketProtocol     socket_protocol = 3;
    optional bytes              query_address = 4;
    optional bytes              response_address = 5;
    optional uint32             query_port = 6;
    optional uint32             response_port = 7;
    optional uint64             query_time_sec = 8;
    optional fixed32            query_time_nsec = 9;
    optional bytes              query_message = 10;
    optional bytes              query_zone = 11;
    optional uint64             response_time_sec = 12;
    optional fixed32            response_time_nsec = 13;
    optional bytes              response_message = 14;
    optional Policy             policy = 15;
}
```

https://github.com/dnstap/dnstap.pb

SURF

Collector

- Go applicatie
- Ontvangt DNSTAP berichten van resolvers
- Berekent hash/versleuteling benodigde data
  - HMAC-SHA256 voor hashes
  - ChaCha20Poly1305 voor versleuteling
- Mogelijkheid voor realtime alerts
- Eenvoudige webinterface voor opzoeken

Database

- Data per rij: 139 bytes
  - ~4 TB voor 90 dagen zonder overhead
- Opslag in een PostgreSQL database
  - Tabelpartities per dag
  - Index op de partities

Prestaties

- DNSTAP non-blocking
- Alleen symmetrische cryptografie
- Simulaties
  - ~5 miljard queries voor 1 miljoen unieke domeinen over 20 dagen, ~1TB data
    - Zoekactie met index op alle tabellen: ~5 sec
    - Zoekactie zonder index in een enkele tabel: ~50 sec
- Ruimte voor optimalisaties indien nodig

Aanvullende maatregelen

- Beperkte toegang tot het systeem
- Audit log
  - Wie heeft wat opgezocht in het systeem?
  - Regelmatige checks
- Gegevens worden een beperkte periode opgeslagen
  - Niet langer dan 90 dagen

Eat your own dog food

- Pilot voor het interne SURF netwerk
  - Alle opvragingen gelogd
  - Real-time checks op bekende kwaadaardige domeinnamen
- Regelmatig alerts voor bedreigingen met een laag niveau
- Een alert met een hoog niveau
  - Dit bleek een gast te zijn

Takeaways

- Pas privacy-by-design toe
  - Welke data is echt noodzakelijk?
  - Hoe kan de data worden opgeslagen met minimale impact op privacy?
- Het is mogelijk om zowel security als privacy te hebben

Vragen?

Joeri de Ruiter

joeri.deruiter@surf.nl

www.surf.nl