EUROPEAN CLOUD SUMMIT 2024

Protecting sensitive data in Azure SQL database

Ivan Vagunin

PhD, MCSM, CISSP

EUROPEAN CLOUD SUMMIT

Microsoft run events

AURUM AvePoint CoreView dox42

EasyLife 365 resco veeam

adesso business. people. technology. Allied Global ASCENT BCC DE CIX devoteam

empowerID FPT Software glueck kanja Jabra GN kaspersky LightningTools

nintex Rencore ShareGate: Spot by NetApp SysCloud Syskit WEBCON LOW-CODE, BUT BETTER.

Use ECS Coins for Swag!

Top 3 win an Atari 2600+

1. Get the app
2. Visit sessions and sponsors, rate sessions
3. Earn ECS Coins
4. Spend ECS Coins

csmmt.eu/app

run events

Who is Ivan?

Ivan Vagunin, Ph.D.

- Senior Software Architect at Tietoevry Create
- Independent security researcher
- MCSM, CISSP
- Amateur gamedev
- Self-proclaimed AI-expert

Microsoft CERTIFIED Solutions Master

CISSP®

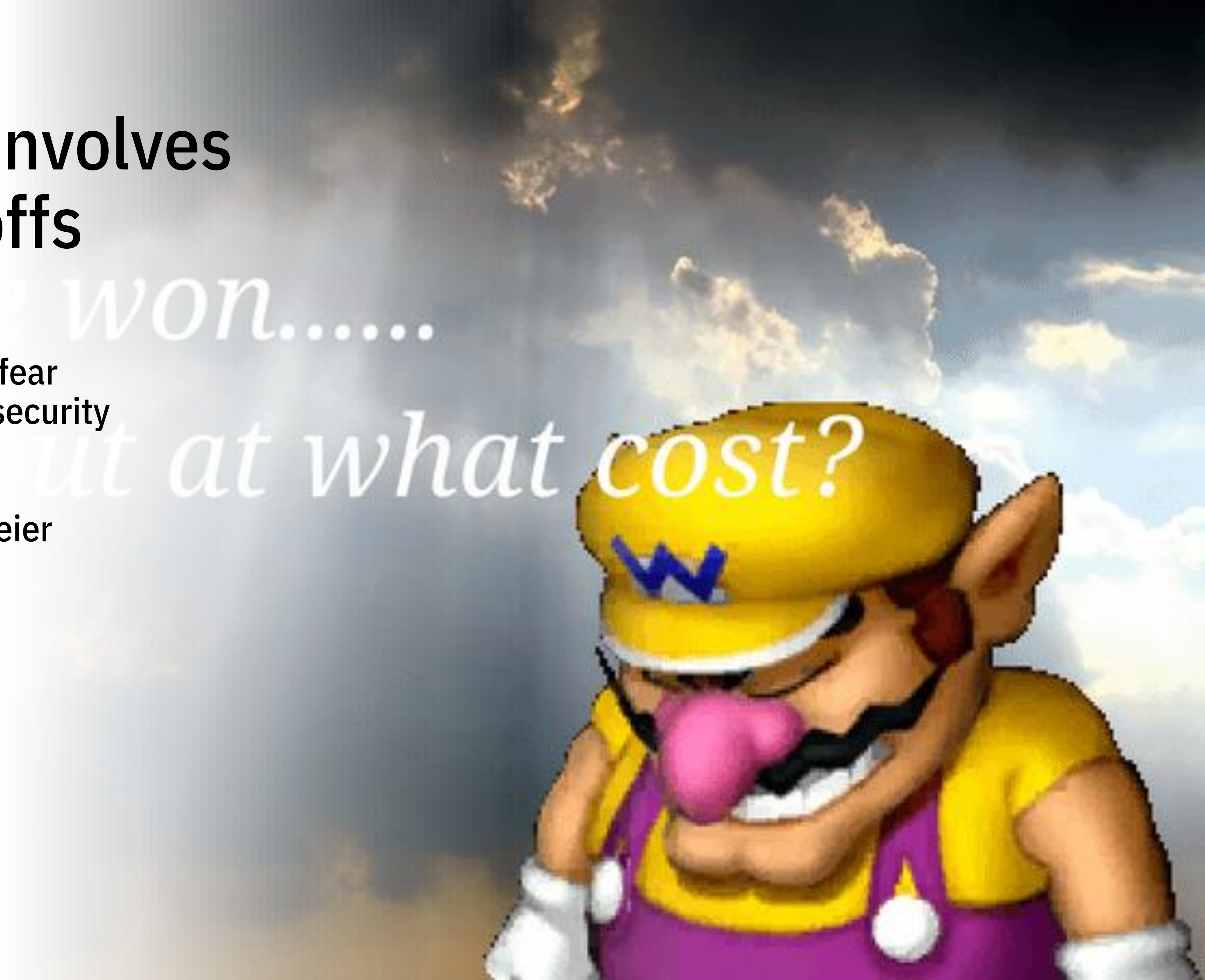
Microsoft CERTIFIED AZURE SOLUTIONS ARCHITECT EXPERT ★★★

Microsoft CERTIFIED CYBERSECURITY ARCHITECT EXPERT ★★★

All security involves trade-offs

“We need to move beyond fear
and start making sensible security
trade-offs”

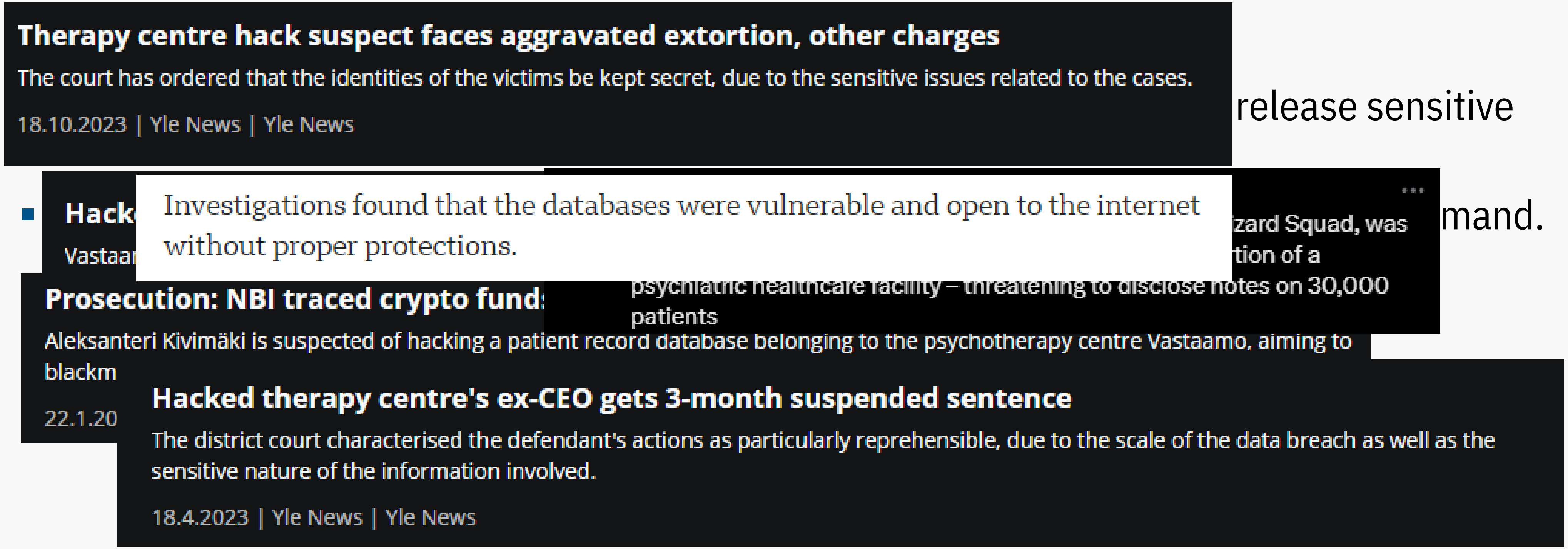
- Bruce Schneier



Remember Vastaamo

release sensitive

- mand.



Therapy centre hack suspect faces aggravated extortion, other charges

The court has ordered that the identities of the victims be kept secret, due to the sensitive issues related to the cases.

18.10.2023 | Yle News | Yle News

Hack

Vastaa

Investigations found that the databases were vulnerable and open to the internet without proper protections.

zard Squad, was tion of a psychiatric healthcare facility – threatening to disclose notes on 30,000 patients

Prosecution: NBI traced crypto fund

Aleksanteri Kivimäki is suspected of hacking a patient record database belonging to the psychotherapy centre Vastaamo, aiming to blackm

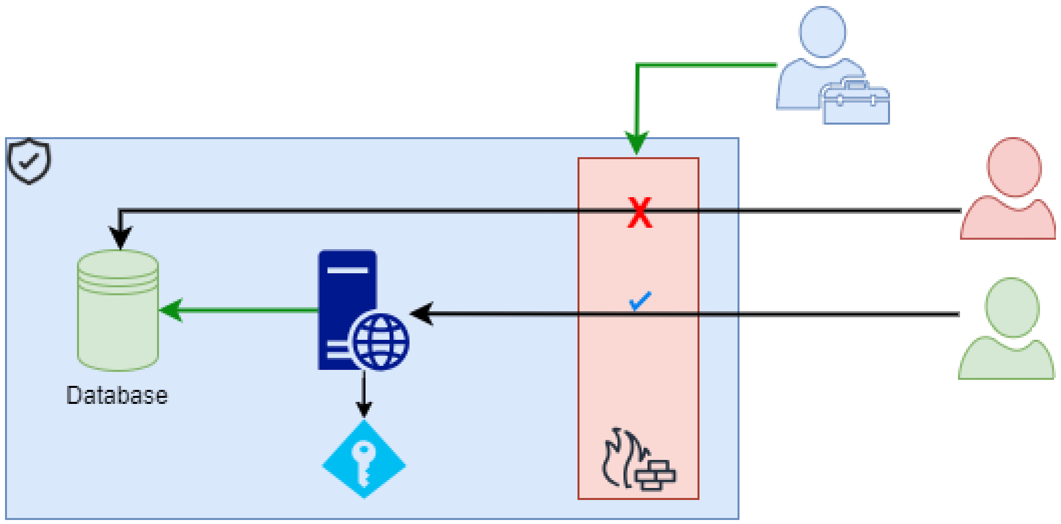
22.1.20

Hacked therapy centre's ex-CEO gets 3-month suspended sentence

The district court characterised the defendant's actions as particularly reprehensible, due to the scale of the data breach as well as the sensitive nature of the information involved.

18.4.2023 | Yle News | Yle News

Vastaamo case



Cloud database threats

Data states

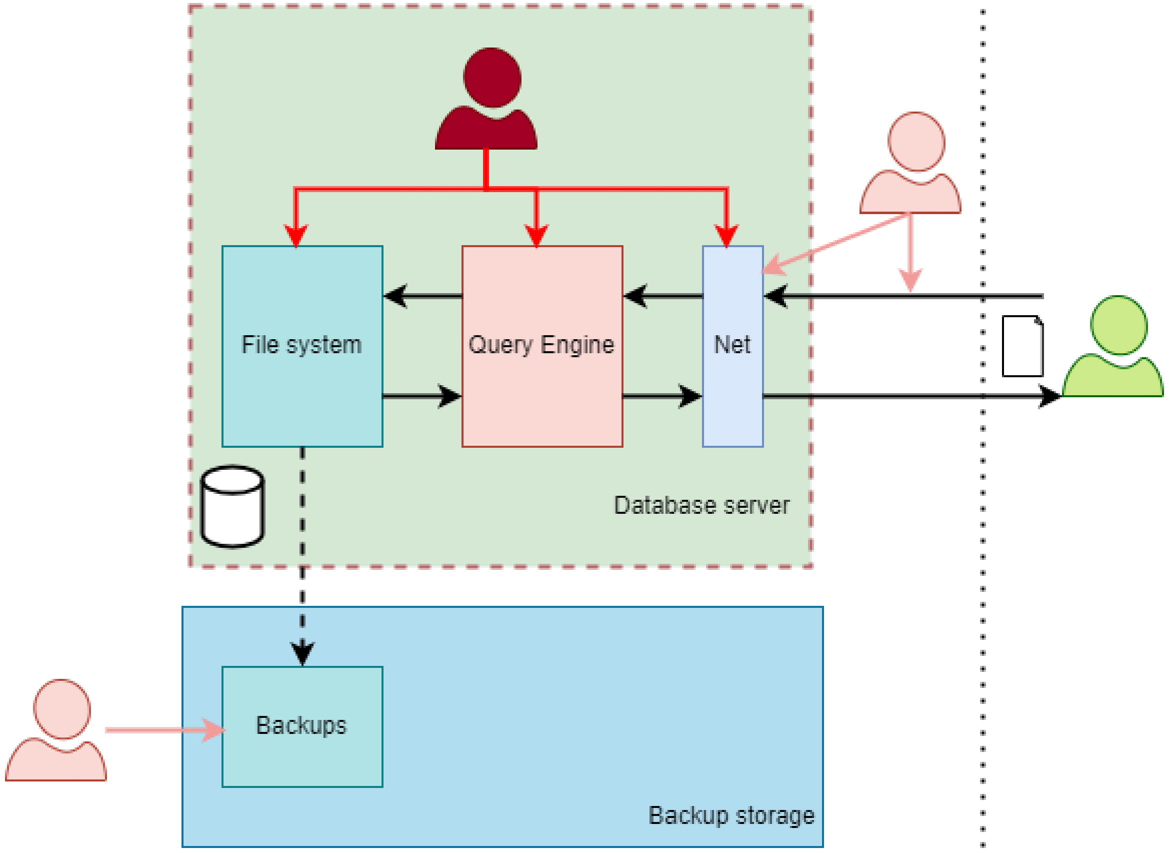
- In-Transit
- At-Rest
- In-Use

Honest but curious (HBC) adversary 🤓

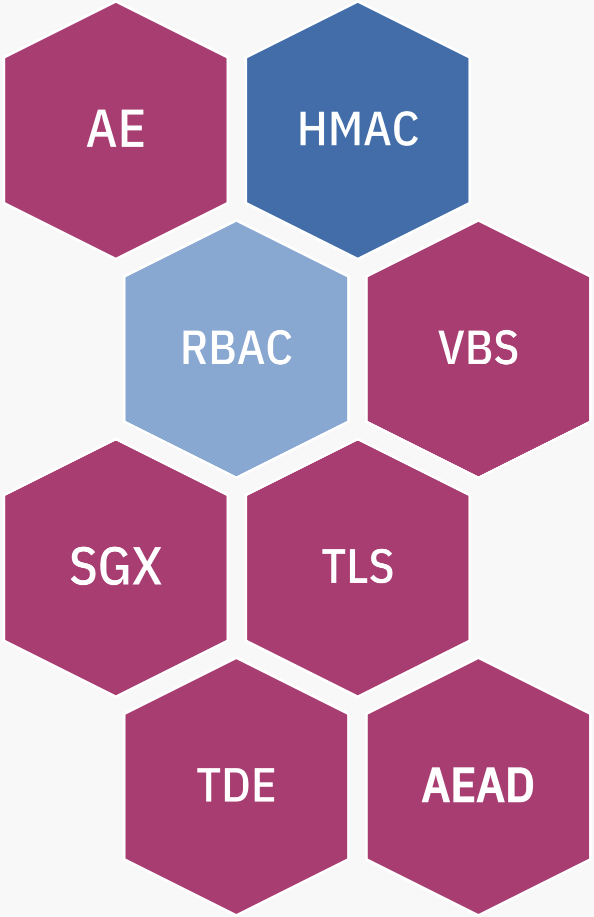
- Observes communication
- Follows protocol

Strong adversary 💪

- Has unbounded power over the SQL Server process
- Can view the contents of the server's memory/disk and all external and internal communication
- Can tamper process (e.g. debugger to SQL.)



Protecting data in Azure SQL database



AE

HMAC

RBAC

VBS

SGX

TLS

TDE

AEAD

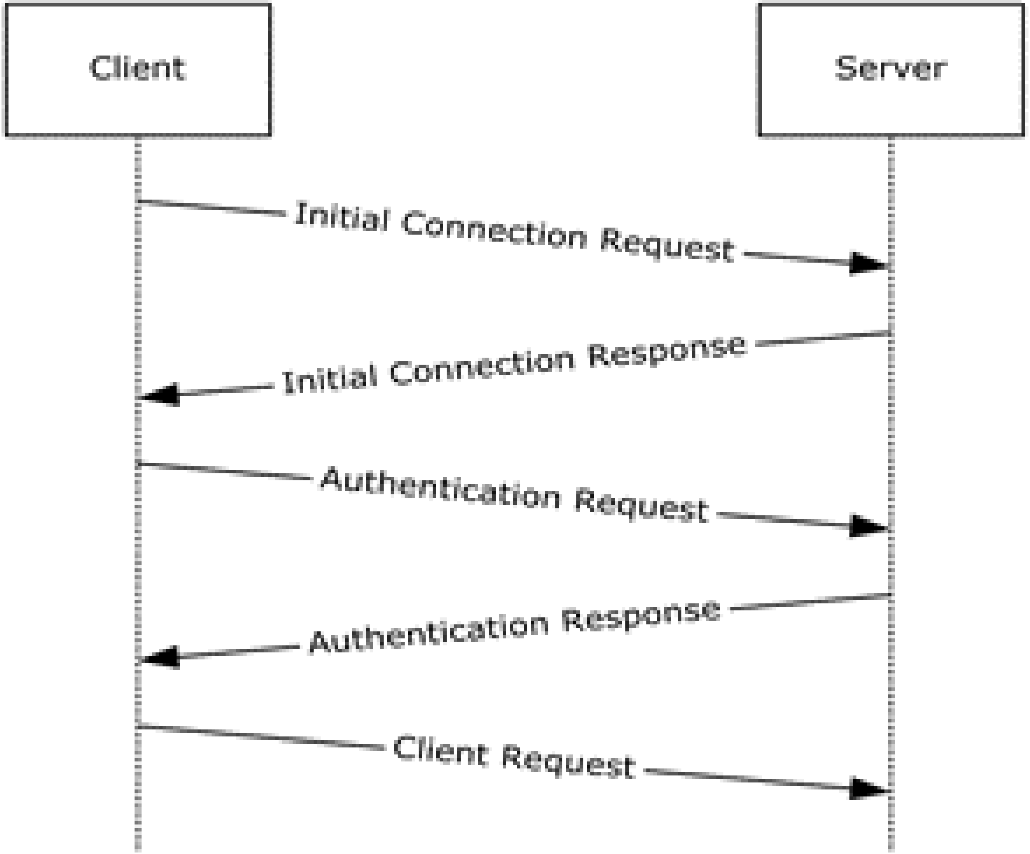
IN-TRANSIT

TDS & TLS

SQL Connection encryption

Tabular Data Stream (TDS)

- In the TDS 7.x version family, encryption is optional and is negotiated and handled in the TDS layer.
- The TDS 8.0 version introduces mandatory encryption that is handled in the lower layer before TDS begins functioning.



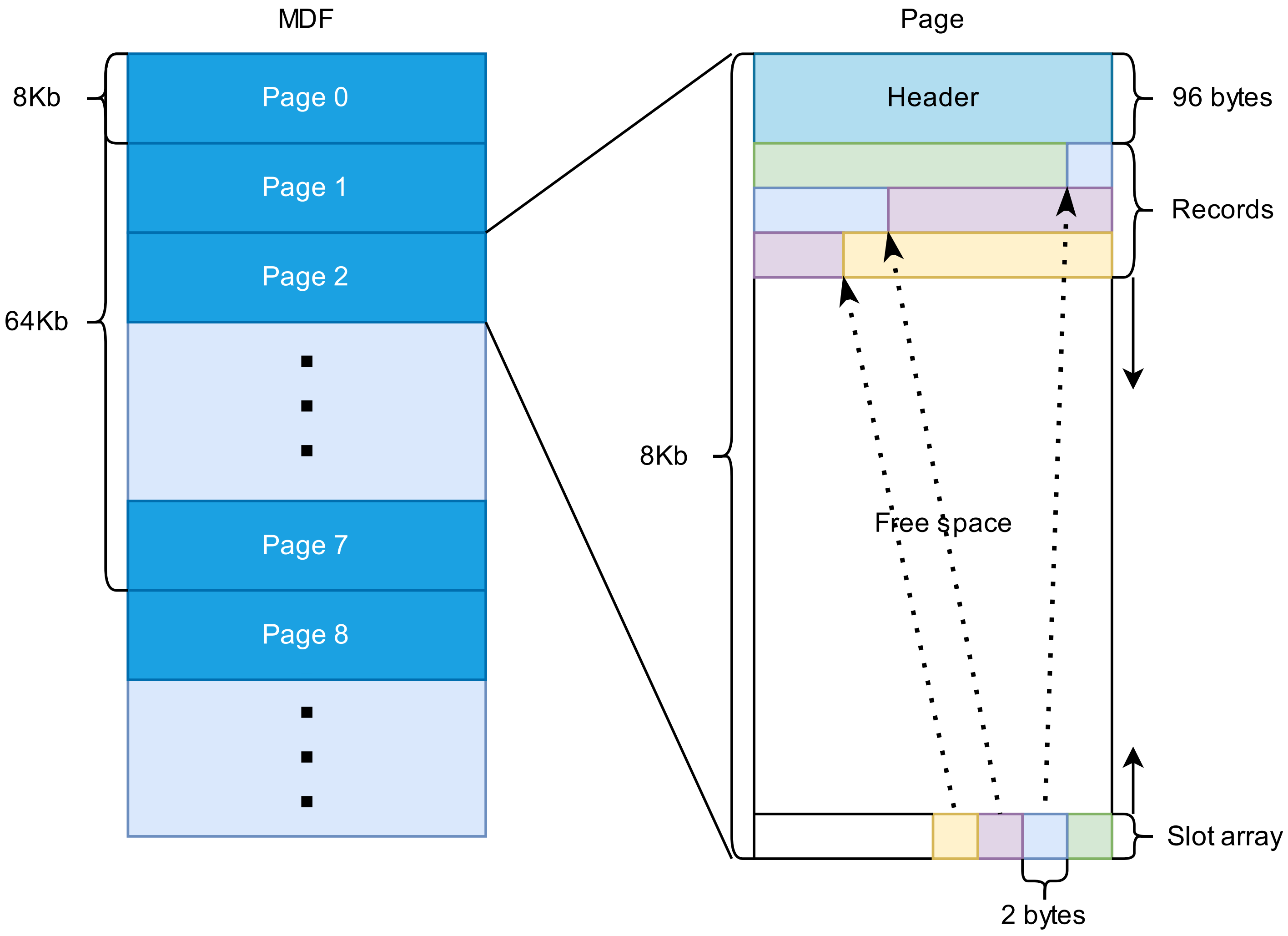
AT-REST

MDF & TDE

Master Data File

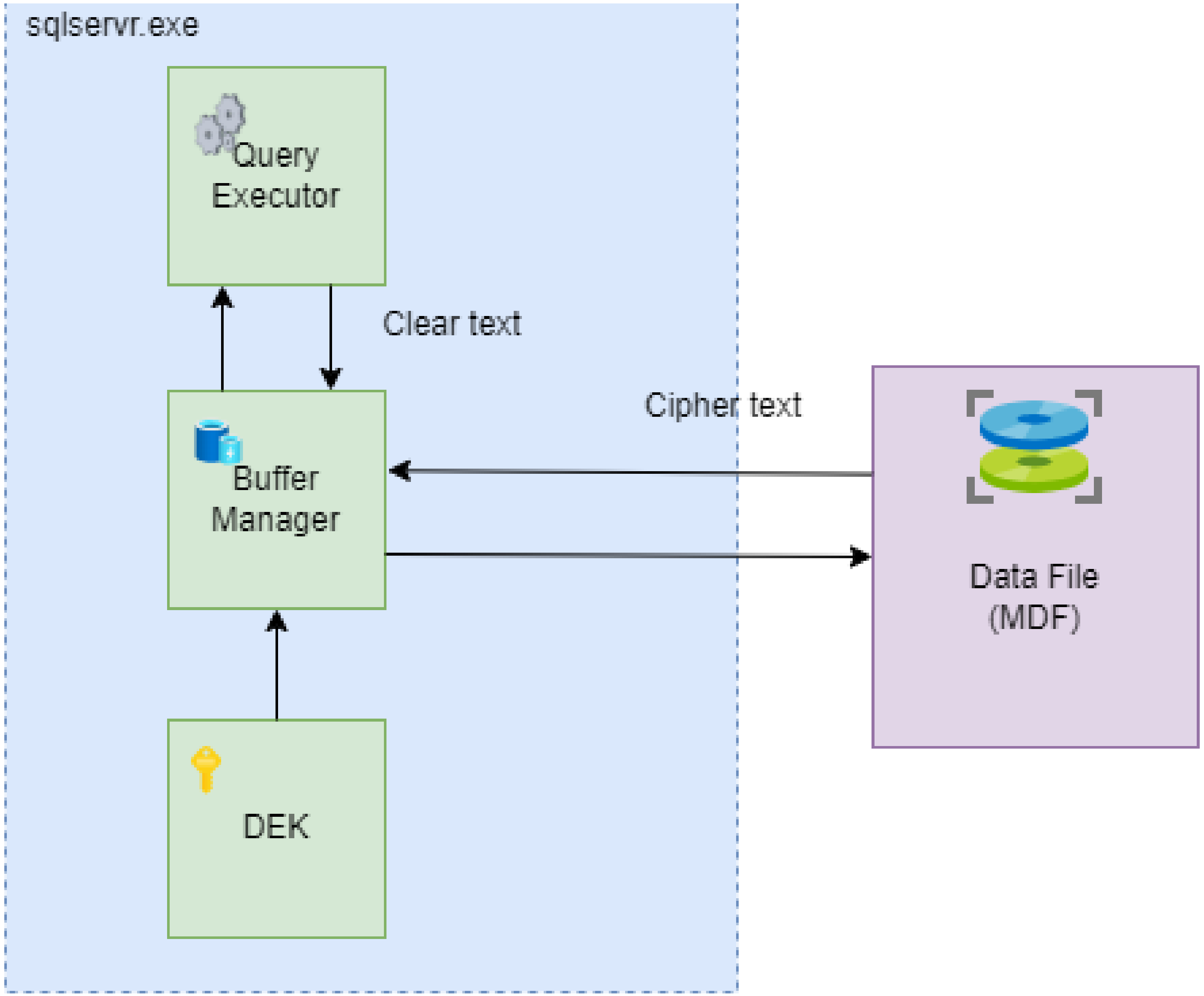
MDF file consists of multiple pages

- The size of a page is 8k bytes
- Header: represents page type, the count of records, free space in the page, and so on.
- Records: vary depending on page type, but generally data records hold actual data associated with a table.
- Slot array: tracks record position.



Transparent Data Encryption

- Once an Azure SQL Database customer enables TDE, keys are automatically created and managed for them
- As of June 2017, Transparent Data Encryption (TDE) is enabled by default on newly created databases. Azure SQL Database supports RSA 2048-bit customer-managed keys in Azure Key Vault.



DEMO

MDF & TDE

IN-PROC

{this space intentionally left blank.}

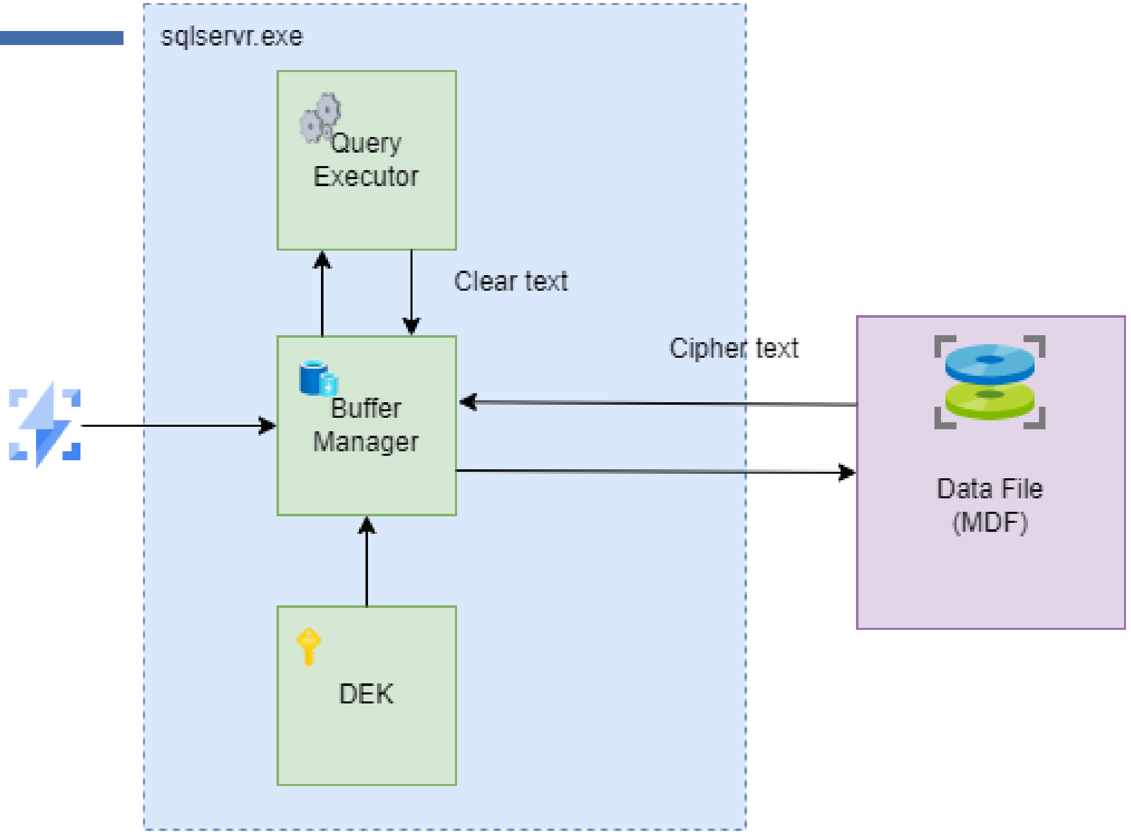
IN-PROC

SQL Server memory architecture

- minimize disk I/O because disk reads and writes
- a buffer pool in memory to hold pages read from the database.

Buffer management

- buffer manager to access and update database pages
- buffer cache (also called the buffer pool), to reduce database file I/O.
  A page remains in the buffer cache until the buffer manager needs the buffer area to read in more data.
- Reading pages
  The Database Engine supports a performance optimization mechanism called read-ahead. Read-ahead anticipates the data and index pages needed to fulfill a query execution plan and brings the pages into the buffer cache before they are actually used by the query.

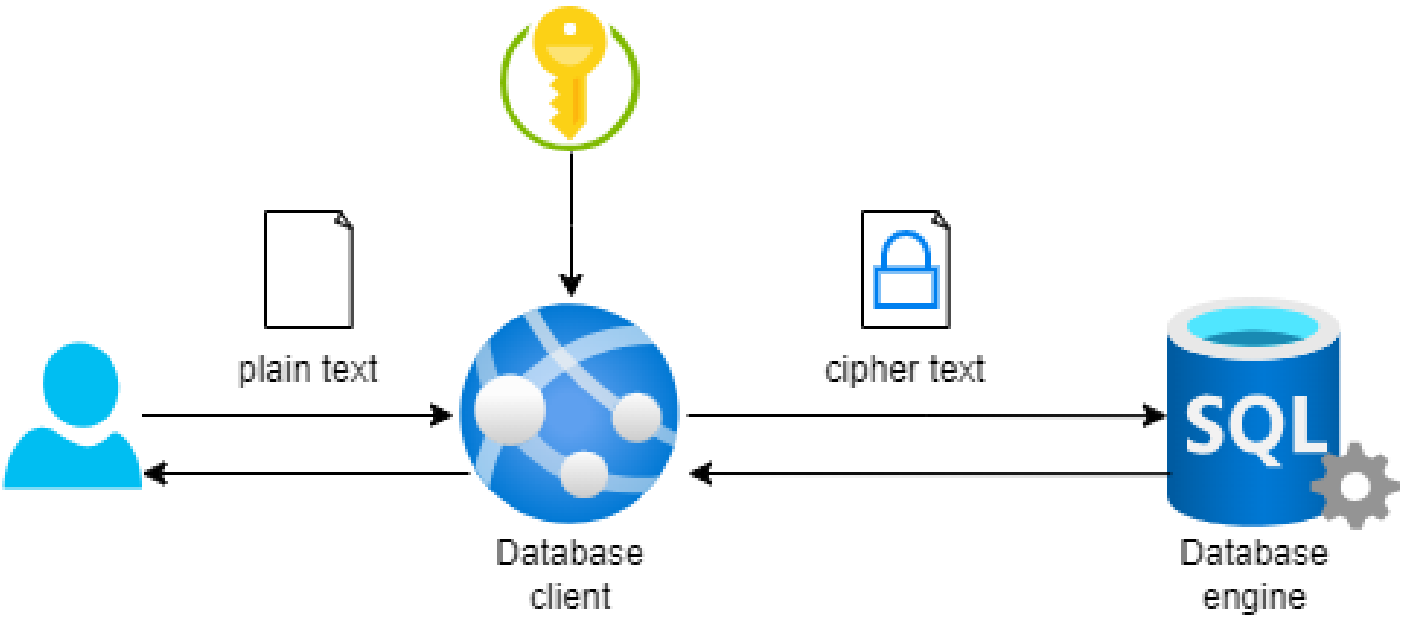


DEMO

READING SERVER MEMORY

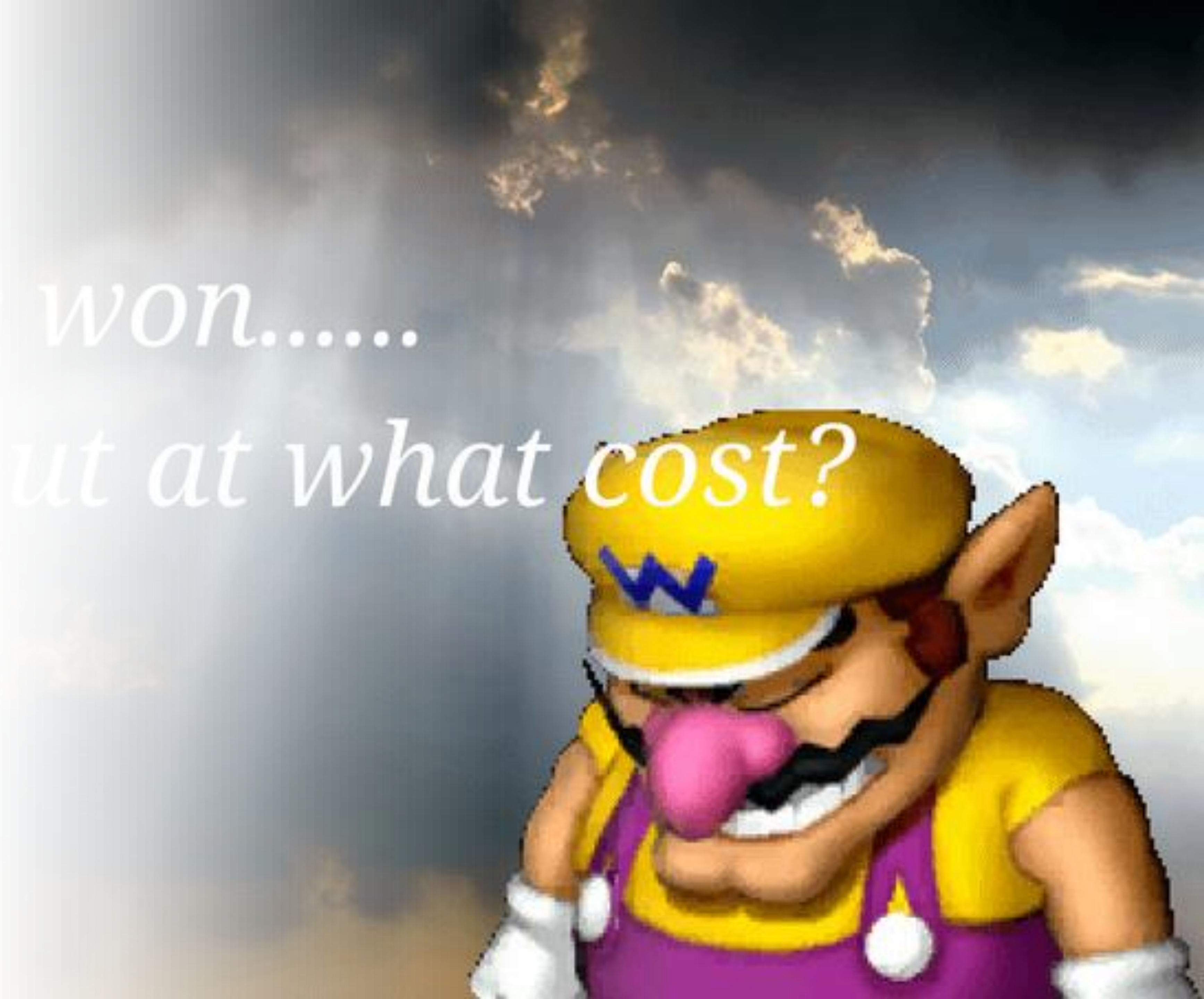
Client-side encryption

- Data is encrypted locally to help ensure its security
- Covers at 3 states of data
- Server can’t decrypt data



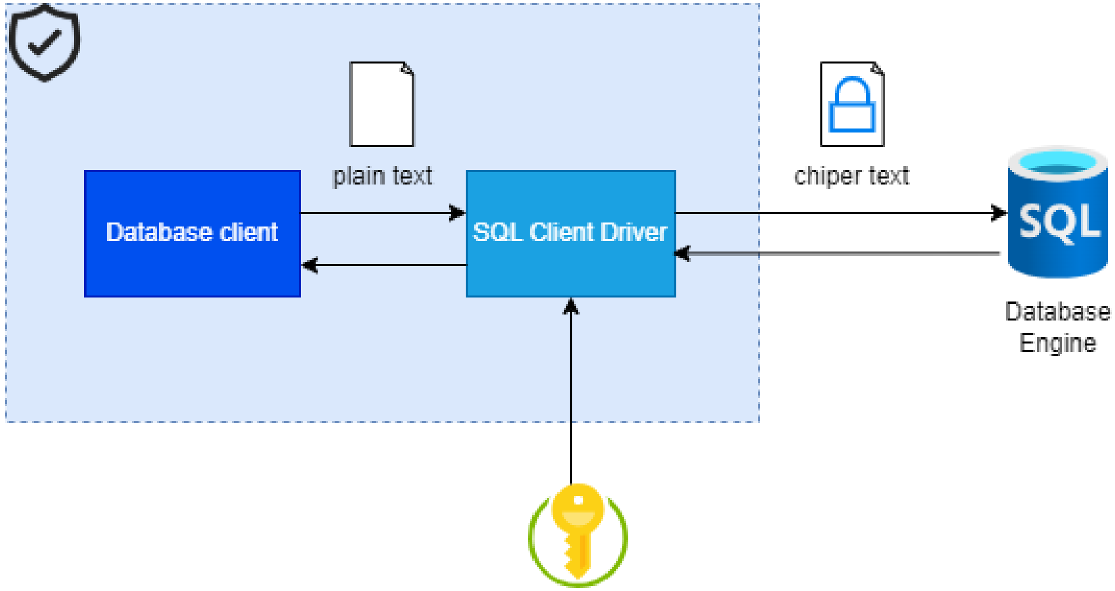
Client-side encryption trade-offs

- Client implications
- Server implications



Always Encrypted

- “Transparent” encryption (for client)
    - Query parametrization
- 2-level key hierarchy
- 2 encryption schemes AEAD
- Column granularity

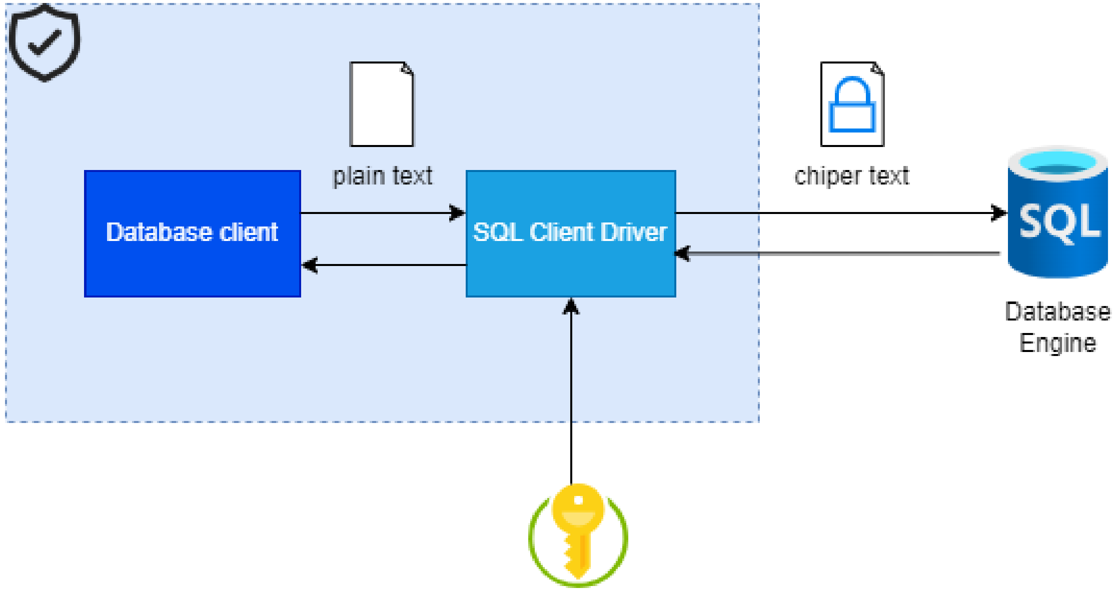


DEMO

ALWAYS ENCRYPTED

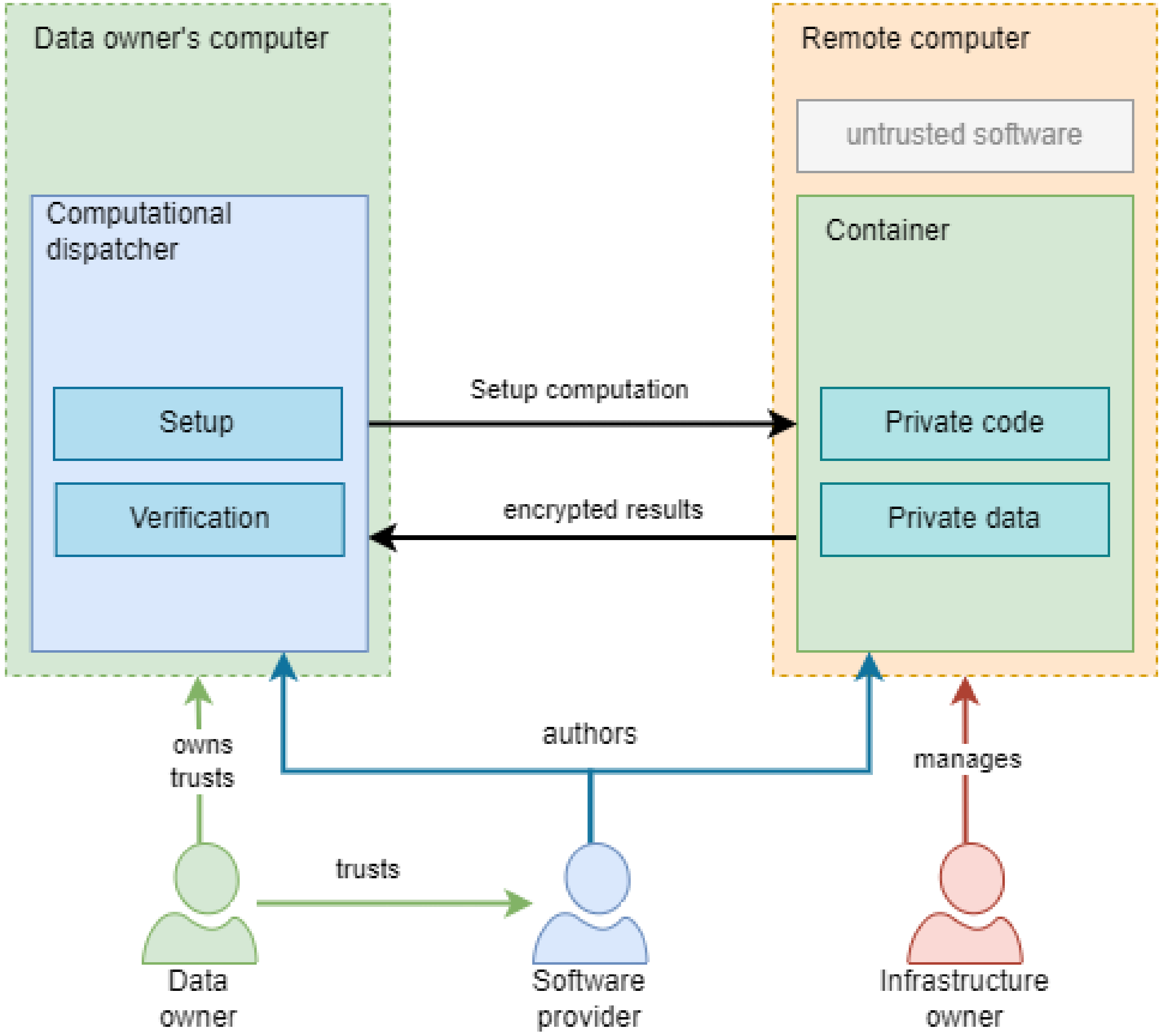
Always Encrypted Trade-offs

- Only equality indexes*
- Double round trip on queries
- Round trip encryption
- Collation limitation: BIN2 or UTF-8 (all case sensitive)



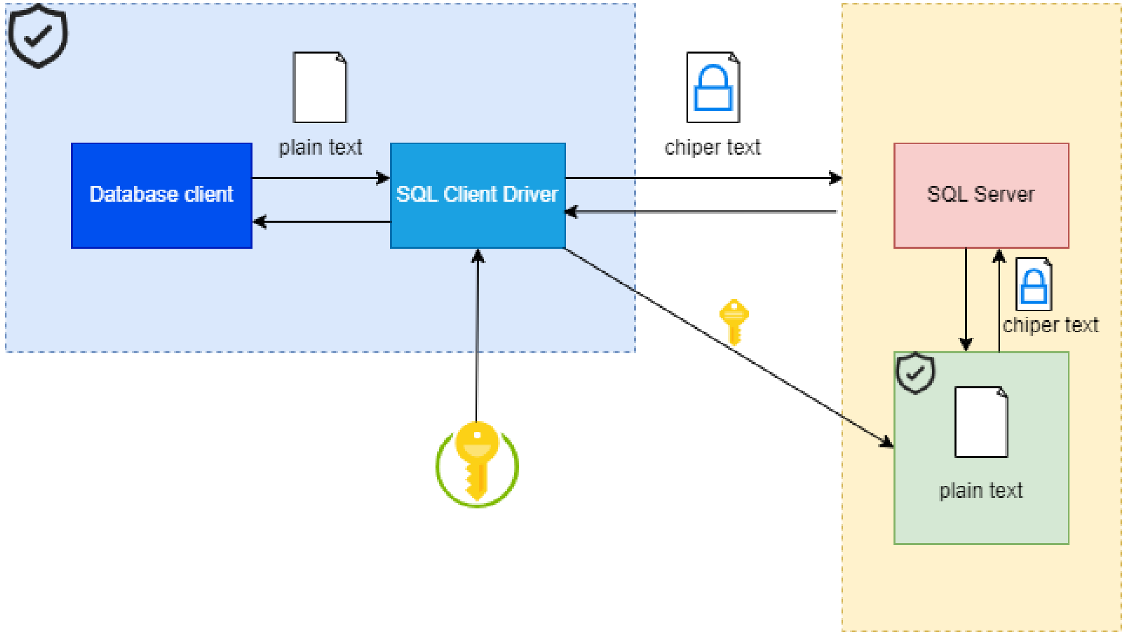
Secure container (enclave)

- Software (Hyper-V/VBS)
  - Relies on Windows hypervisor and doesn't require any special hardware
- Hardware-based (Intel SGX)
  - Intel SGX protects data actively being used in the processor and memory

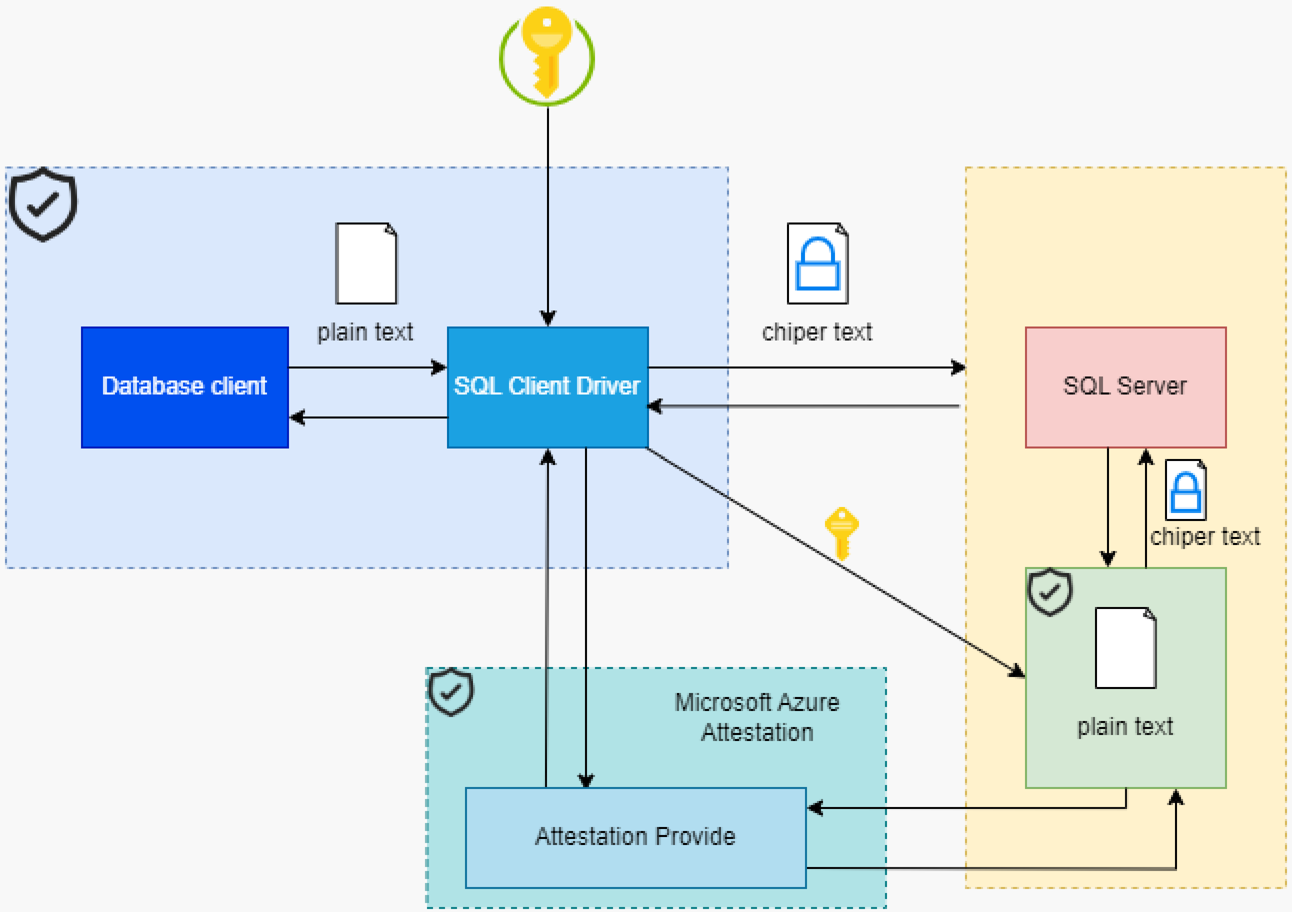


Always encrypted with secure enclaves

- Secure computations in enclave
- Rich queries
- In-place encryption



Attestation



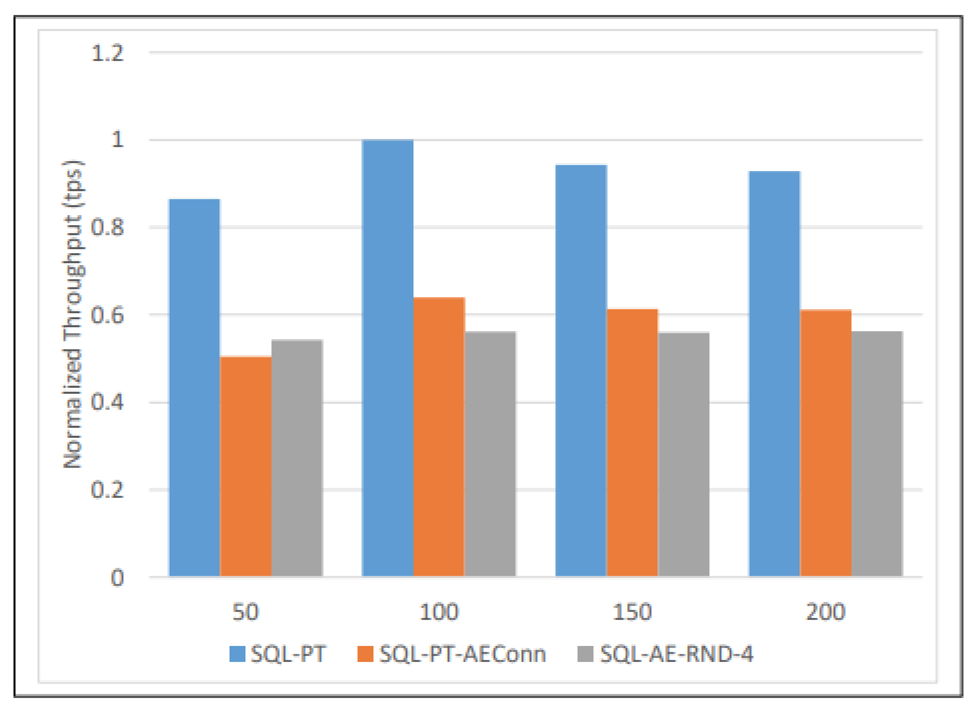
Secure enclave threats

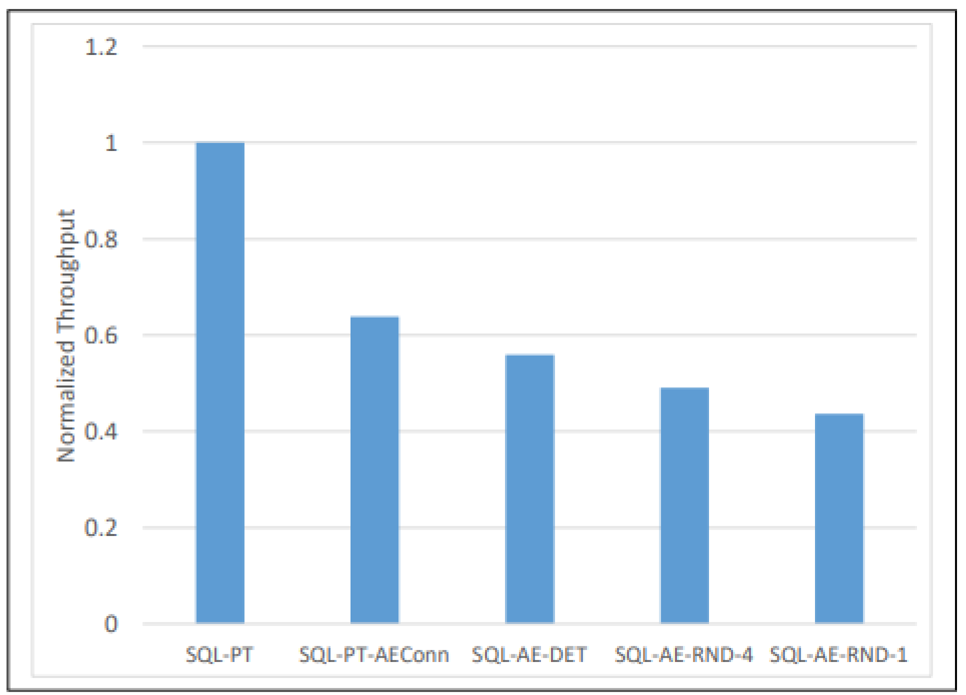
- Query confidentiality
- Index confidentiality
- DDL (Initial encryption/key rotation)
- Metadata related attacks
  - Malicious metadata
  - Malicious key
- Session attacks (replay CEK to enclave)

Operation	Exposure
Comparison (DET)	Frequency distribution over values
Comparison (RND)	Ordering over values
LIKE predicate using scans	Unknown predicate over values
LIKE predicate using an index (i.e. prefix matches)	Ordering over values plus some information about proximity
DDL to encrypt data	Limited access to encryption oracle only with client authorization

Performance

- Calls to AE are expensive
  - Double round-trip (disable AE operations if not needed)





Price

TLS, TDE

- No additional cost

Always encrypted

- No direct additional cost

Always encrypted with secure enclaves

- VBS (test environments)
    - No additional cost
- Production environment
    - SGX-enclaves only available on DC-series
    - ~19.5 EUR/Day (DTU ~6 EUR/Day)

What encryption to use

ENCRYPTION	USE-CASE
TLS	ALWAYS
TDE	ALWAYS
Always Encrypted Deterministic	Highly sensitive data, no wildcard search, non-repeating large set
Always Encrypted Randomized	Highly sensitive data, no search
Always Encrypted with Secure Enclaved Deterministic Scheme	Never?
Always Encrypted with Secure Enclaved Randomized Scheme	Highly sensitive data, wildcard search

Info

- https://aka.ms/always-encrypted-enclaves-docs
- https://aka.ma/ae-paper-sigmod-20
- https://github.com/ycherkes/OrcaSql
- https://eprint.iacr.org/2016/086
- https://ieeexplore.ieee.org/document/7113304
- https://web.cs.ucdavis.edu/~franklin/ecs228/2013/popa_etal_sosp_2011.pdf
- https://www.kazamiya.net/mssql_4n6-01
- https://github.com/ivanvagunin

THANK YOU,
YOU ARE AWESOME ❤️

PLEASE RATE THIS SESSION IN THE MOBILE APP.

{this space intentionally left blank.}