# Securing Data in the Cloud

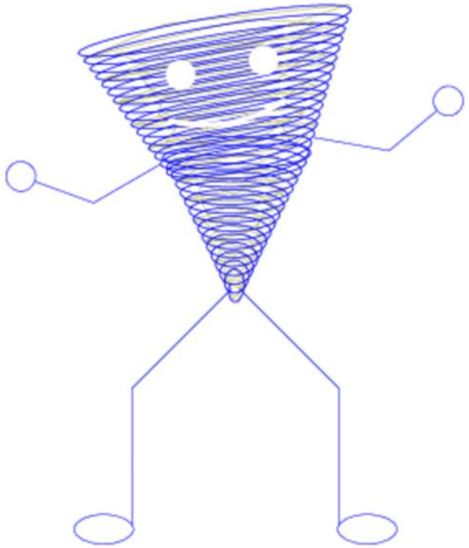
## Making the Most of Trusted Hardware

Nick Felts, Computer Systems Researcher
ngfelts@uwe.nsa.gov

# Today's Presentation

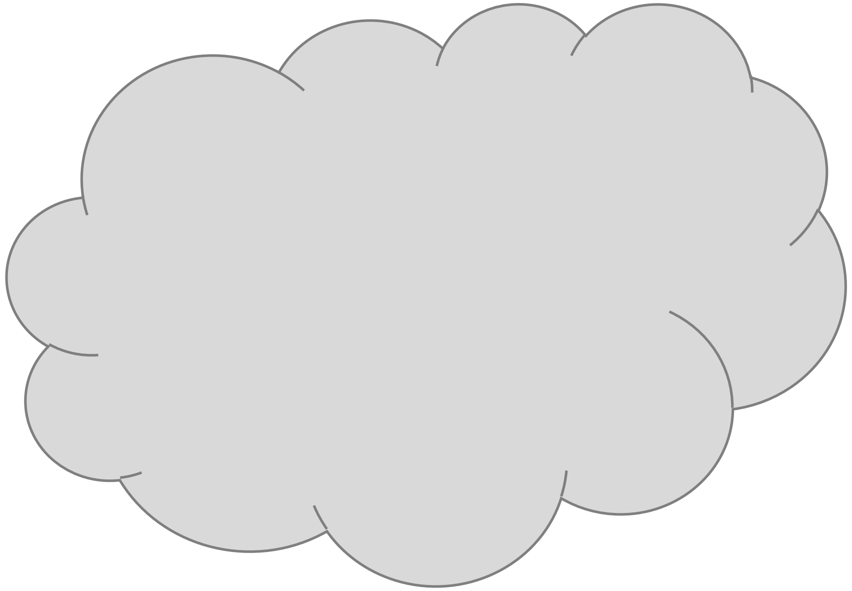
- What is the cloud?
- How do we secure data within the cloud?
- How can we move towards securing data-in-use?
- How can trusted hardware help?

- Not tied to specific vendors



# Large Cloud Services

- Gartner Magic Quadrant [1]
- Amazon Web Services
  - Amazon Key Management Service
- Microsoft Azure
  - Azure Key Vault
- Google Cloud Platform
  - Google Cloud Key Management Service



- Each provides its own Identity and Access Management (IAM)

# The Cloud

https://upload.wikimedia.org/wikipedia/commons/2/27/IBM_360-44.5.jpg

- What do we mean by the cloud?

- Services offered
  - Infrastructure
  - Platform
  - Containers
  - Software
  - Functions

- Major providers vs smaller installations

# Trusted Hardware

- Hardware Security Modules (HSMs)
- Trusted-Execution Environments (TEEs)
  - Enclaves
  - Examples:
    - Intel's Software Guard Extensions (SGX)
    - ARM TrustZone
    - AMD, IBM, and RISC-V also have offerings
- Hardware Root-of-Trust
  - Trusted Platform Modules (TPMs)
  - Other "Security Chips" (e.g., Apple T2)
  - Microsoft Pluton (recently announced)

# Assumptions for Enclaves

- Protected secrets
  - Keys never exposed outside trusted hardware
  - Data within enclave always encrypted
- Protected execution
  - Prevents unauthorized access by external entities
    - e.g. operating system or system admins
  - Attestation mechanisms available
- Usability
  - Minimal impact to performance
  - General purpose
  - Widely available

# Trusted Hardware from Large Services

- Hardware Security Module access
- Trusted-Execution Environments
  - Azure offers SGX access
  - Google is developing Asylo
  - Amazon offers Nitro for EC2

# Introduction to Data Protection at Scale

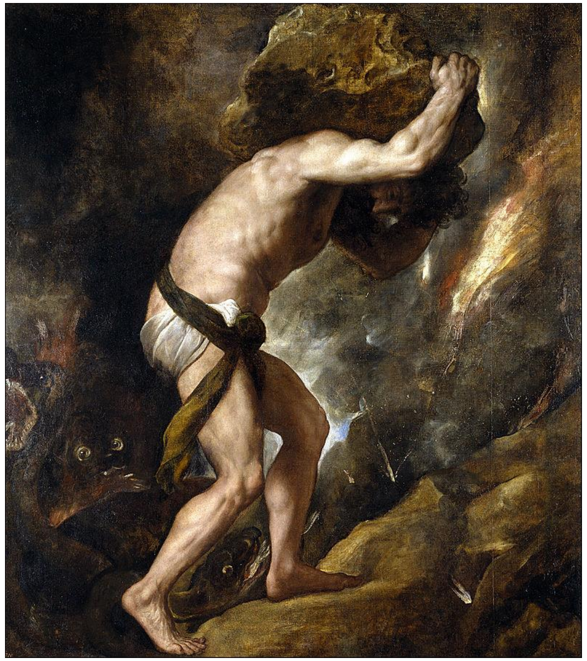
- How useful is data encryption [2]?
  - Useful for deleting data
  - "Default key policies do not offer much security beyond default access controls"
  - It is difficult to move beyond storage layer encryption

- Data should be accessible only at the discretion of the owner [3].
  - Data encryption within clouds should protect its confidentiality
  - Access should be restricted to the owner and others the owner permits

# The Problem

- Meaningful encryption, at scale, is hard.

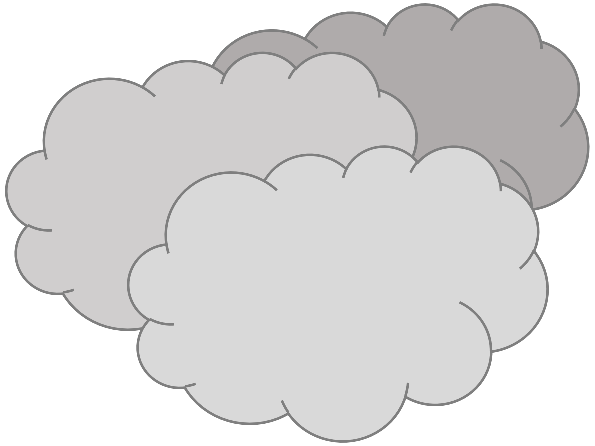
- Several reasons:
  - Managing keys is hard
  - Untrusted system administrators
  - Additional cryptographic processing overhead
  - Threats aren't always clearly stated or understood
  - Legal protections exist – security can become a series of checkboxes



https://commons.wikimedia.org/wiki/File:Punishment_sisyph.jpg

# Private Clouds

- Commodity Hardware
  - Trusted hardware
    - TPM
    - TEEs
  - HSM support may not exist
  - Specially designed hardware not expected
  - Access to hardware more realistic
- Commercial Cloud Providers offer similar services
  - SaaS



# Data-at-Rest Encryption at Scale

- Objects encrypted in storage
- Distributed databases encrypted at rest
  - Great example of data-at-rest
    - Often part of a cloud's offerings
    - A natural place to store heterogeneous data
  - Many layers to do this
    - Local disk
    - Distributed filesystem
    - Within the database itself

# Security vs Performance

- Owners retain keys
  - Lose power of cloud computing
  - Must manage keys themselves
    - Storage, use, revocation, etc
  - Most secure option [4]
  - Available to all users independent of cloud provider

- Cloud services manage keys
  - Leverage cloud computing resources
  - Key exposure
    - To software in the cloud
    - To system administrators
  - Key management flexibility
    - Owners/users can manage keys
    - Cloud can manage keys

# Key Management

- Difficult to meaningfully automate within a cloud setting
  - Who controls/stores the key?
  - Where is the key stored?
    - HSM/TPM
    - Local file system
    - Distributed file system
    - Key server
- Trusted hardware
  - Can commodity trusted hardware provide any assistance?

# Key Management (cont'd)

- Trusted Platform Modules
  - Offer measurement of system
  - Only provide access to keys if measurements pass
- Trusted-Execution Environment
  - Allow for execution of trusted software
  - Minimize or prevent exposure to administrators and other users
  - Killer App?
    - Azure using SGX for AI processing on hospital data sets [5]

# Data-in-Use Protection at Scale

- What is the best we can do for now?
  - Homomorphic encryption vs trusted hardware

- Not yet available as a general solution for at-scale cloud processing

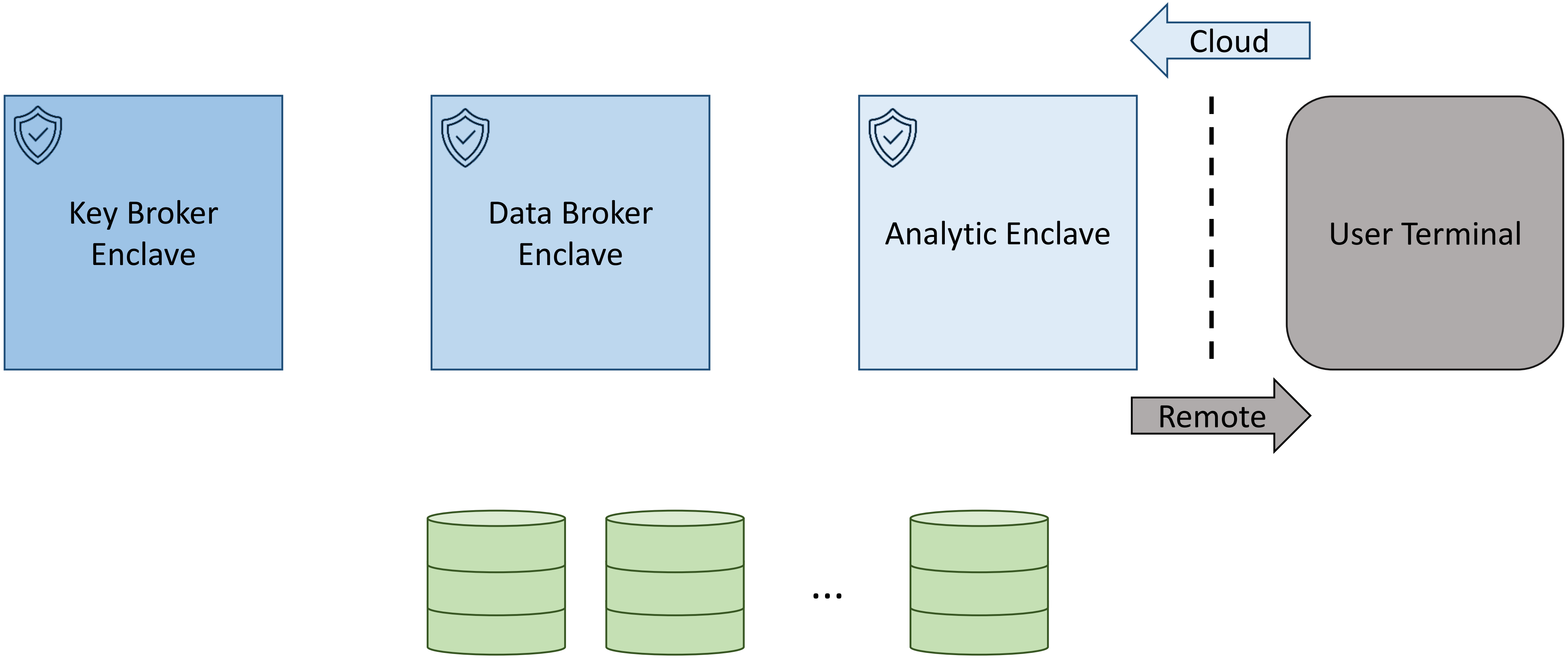
- TEEs are becoming more capable

- TEEs are a practical option for some problems

# Keys-in-Use Protection at Scale

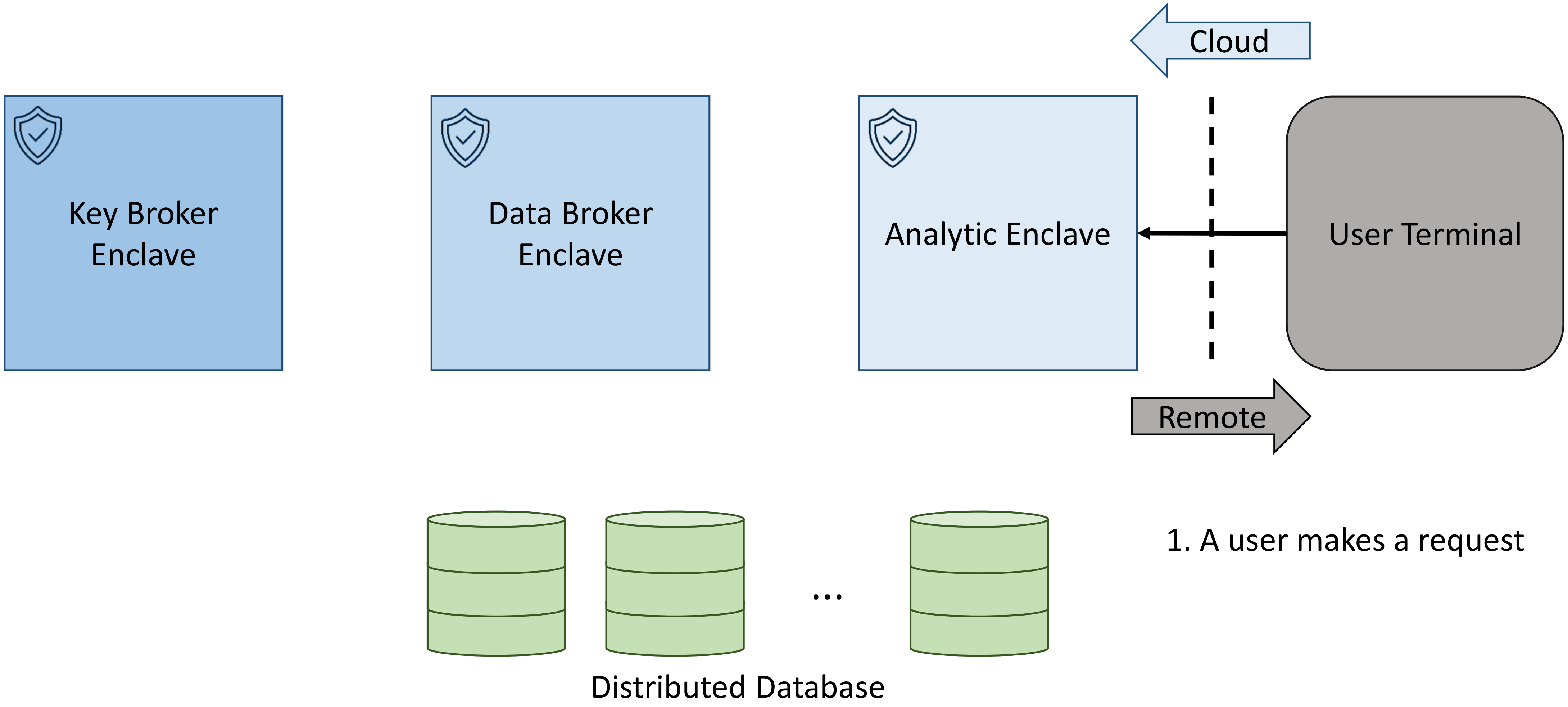
- Trusted hardware
- Limit access to users and tools
  - Integrated into IAM
- Can be done by design
  - Toggled on/off
- Data still exposed, but exposure is limited
- Performance hits
  - Minimized by focusing on TEEs instead of HSMs
  - TEEs can be used to cache keys closer to the software (database)

# Bird's-eye View of Data-in-Use Protections



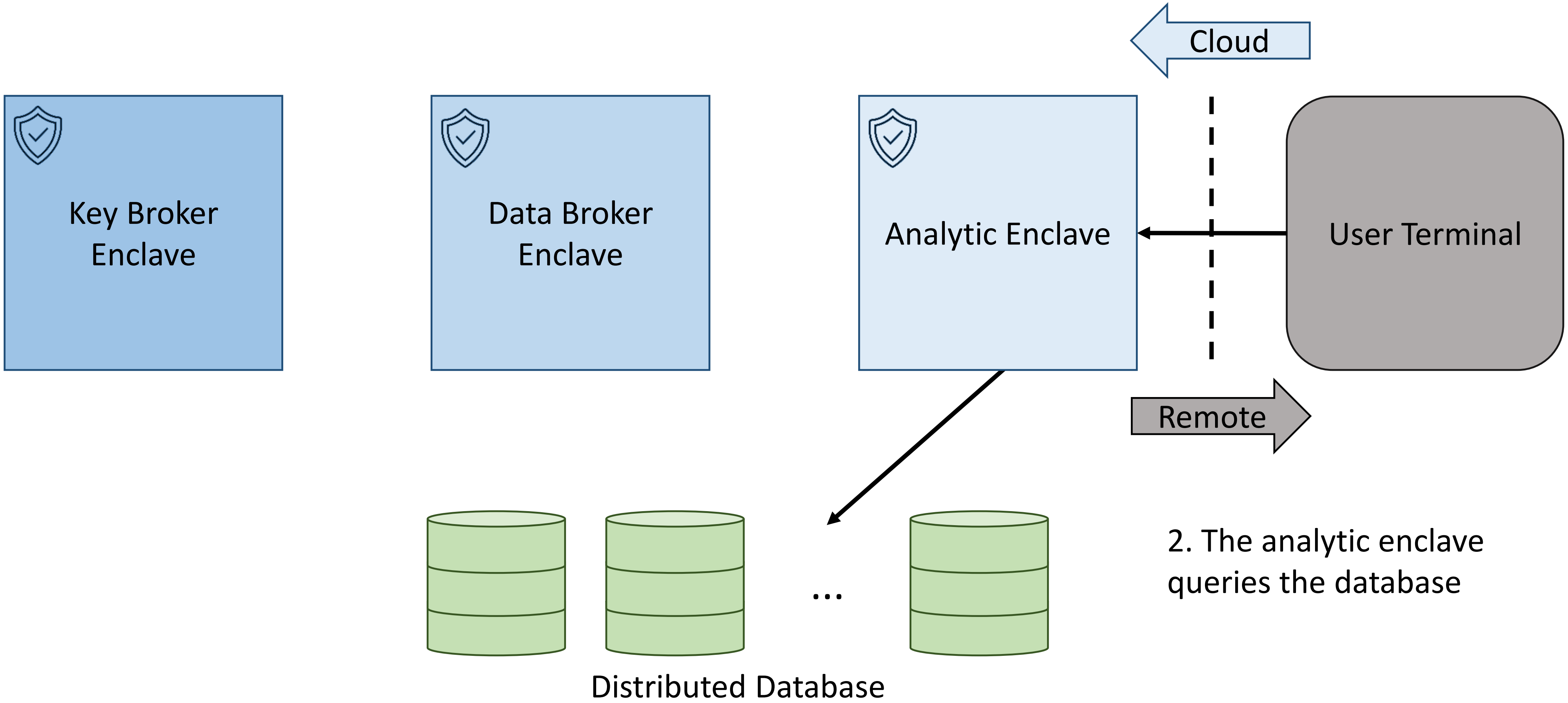
Distributed Database

# Bird's-eye View of Data-in-Use Protections



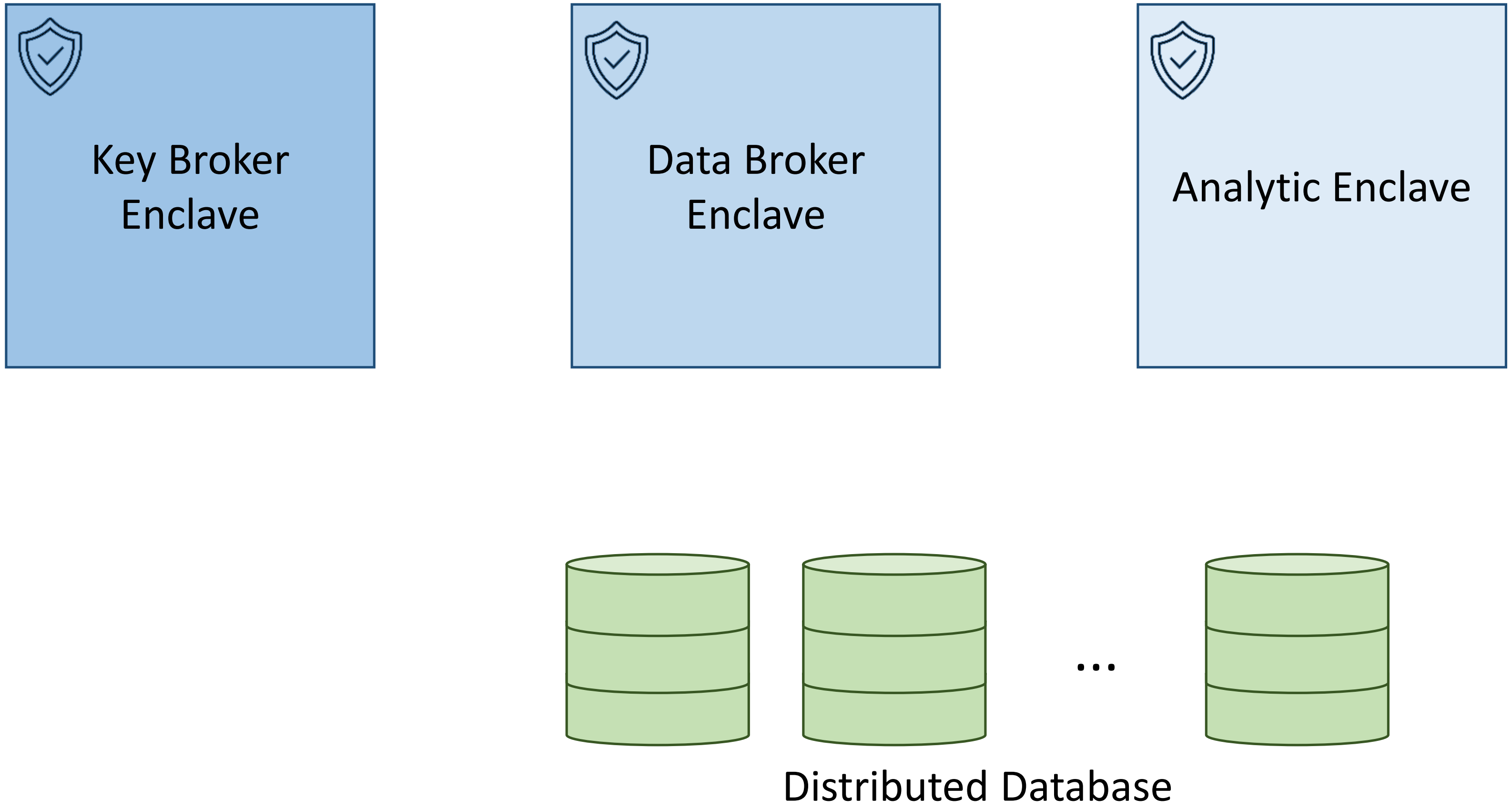
1. A user makes a request

# Bird's-eye View of Data-in-Use Protections

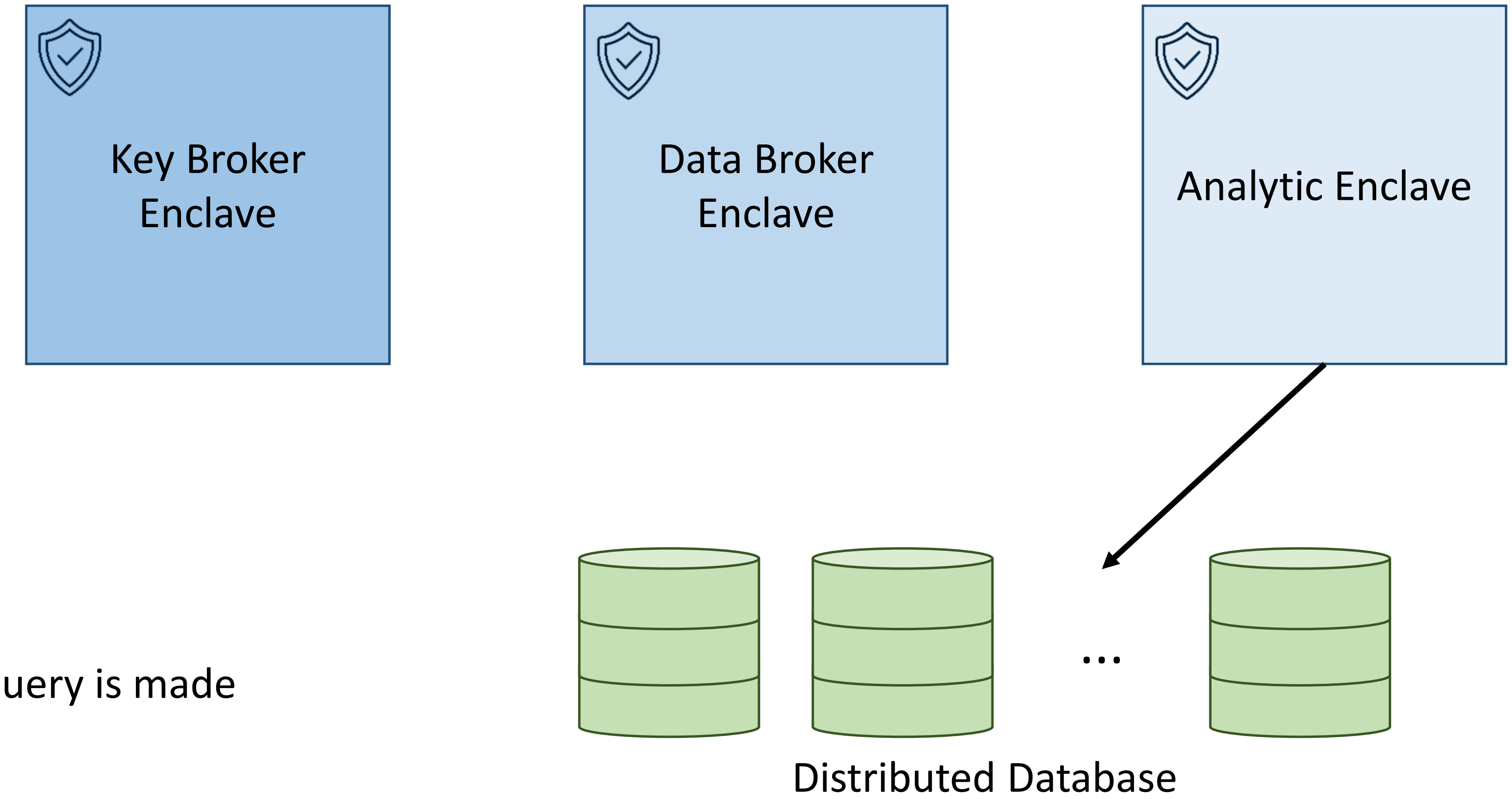


2. The analytic enclave queries the database

# Database Queries

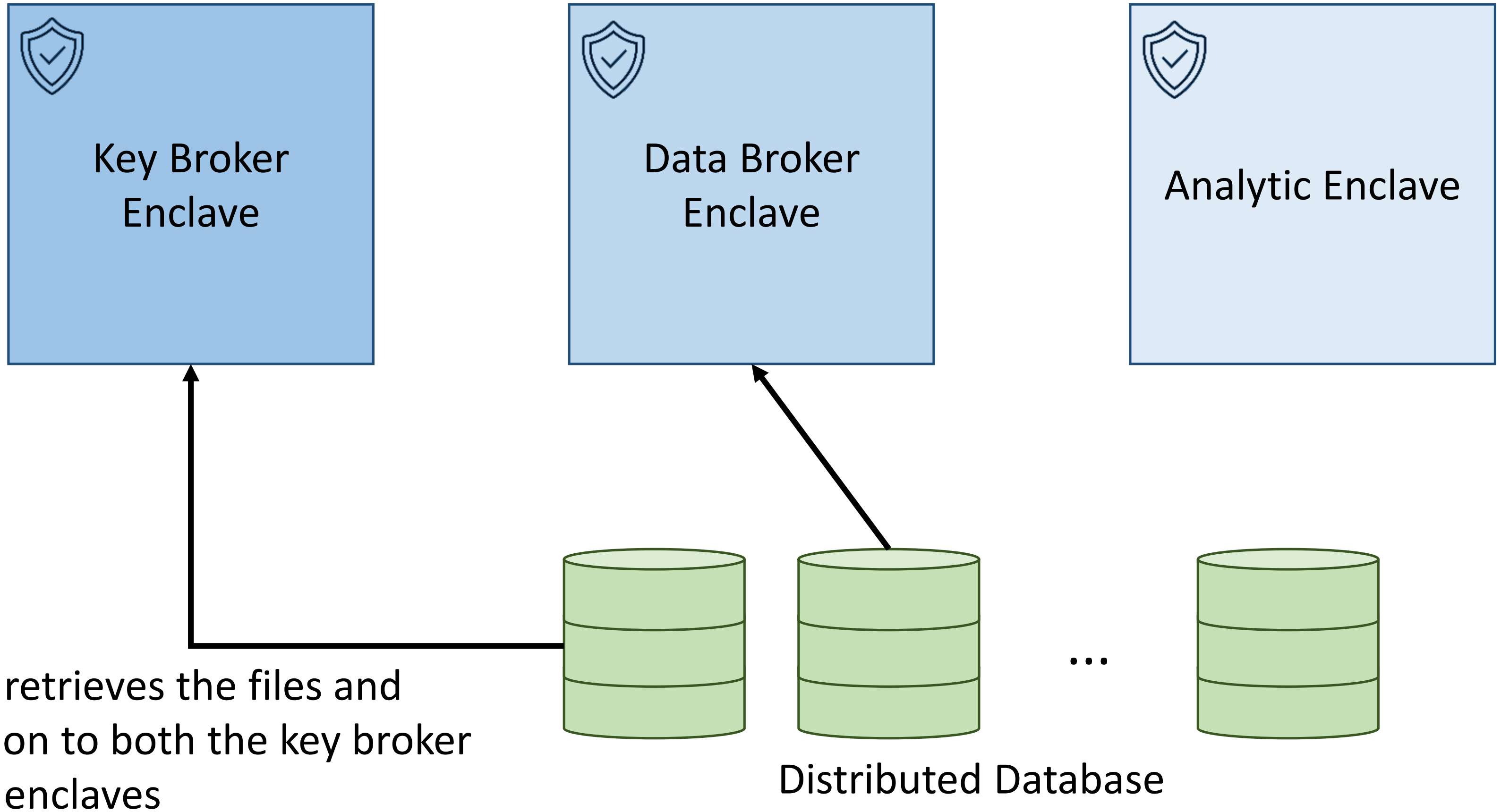


# Database Queries



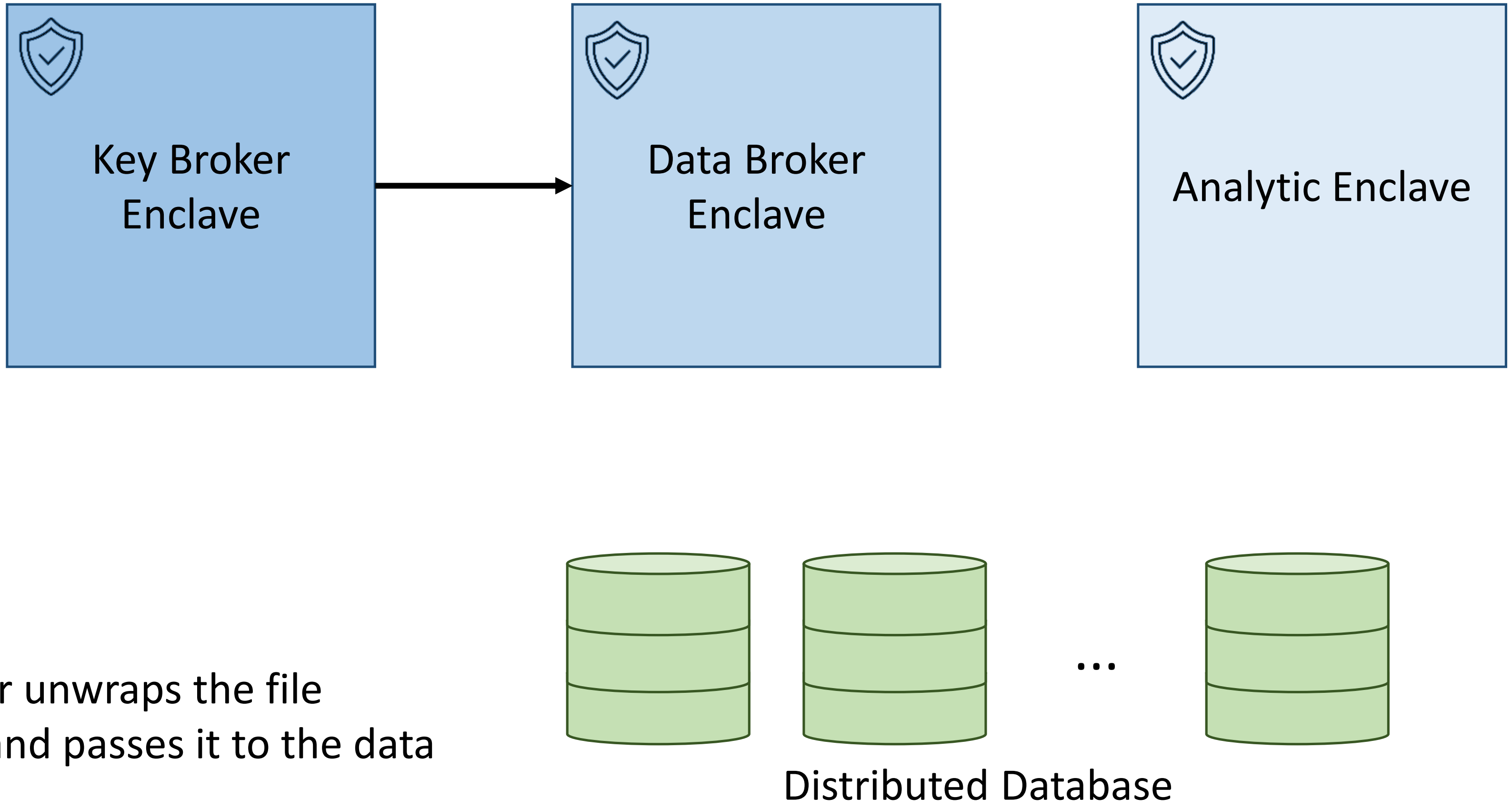
a. A database query is made

# Database Queries



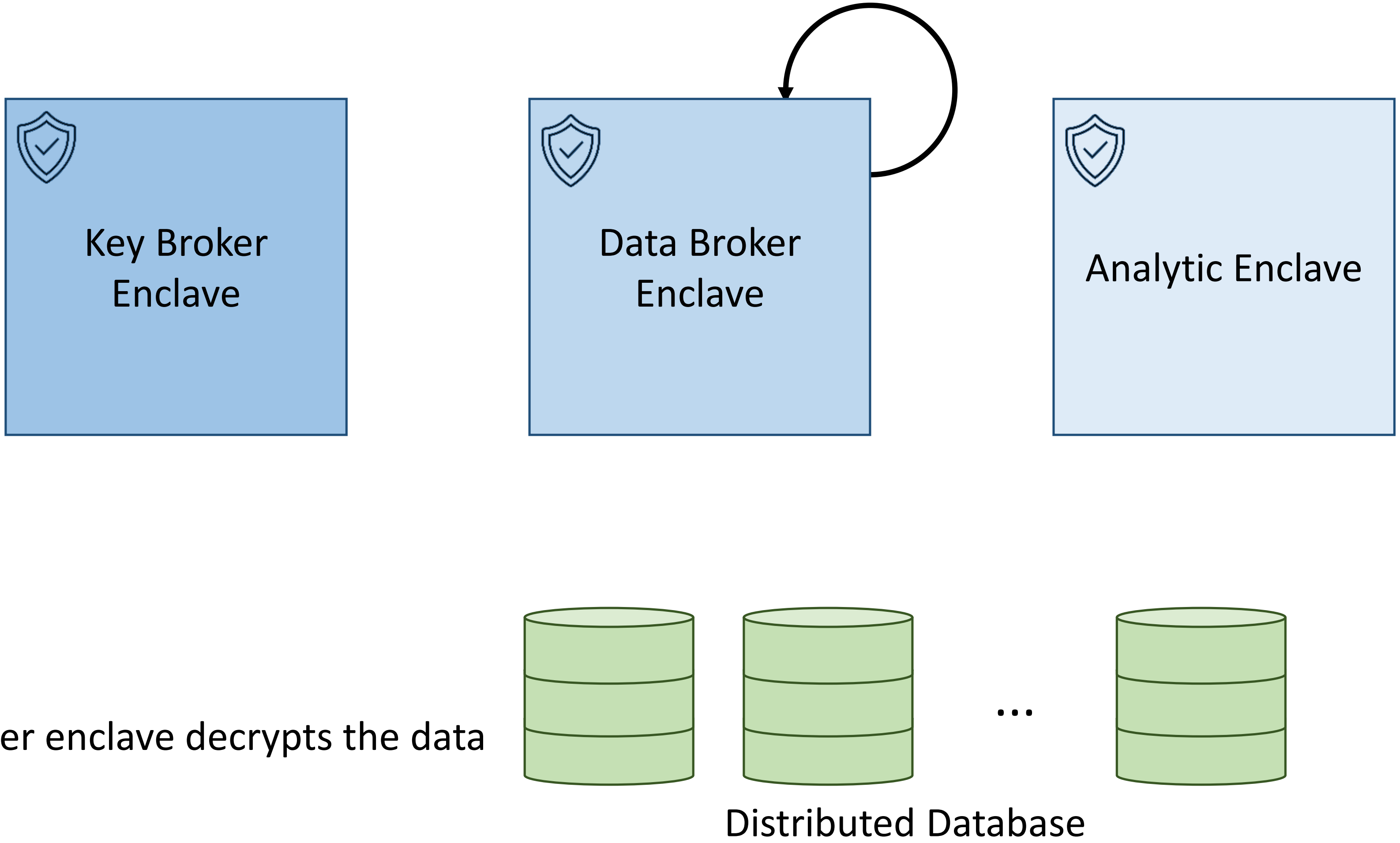
b. The database retrieves the files and passes information to both the key broker and data broker enclaves

# Database Queries



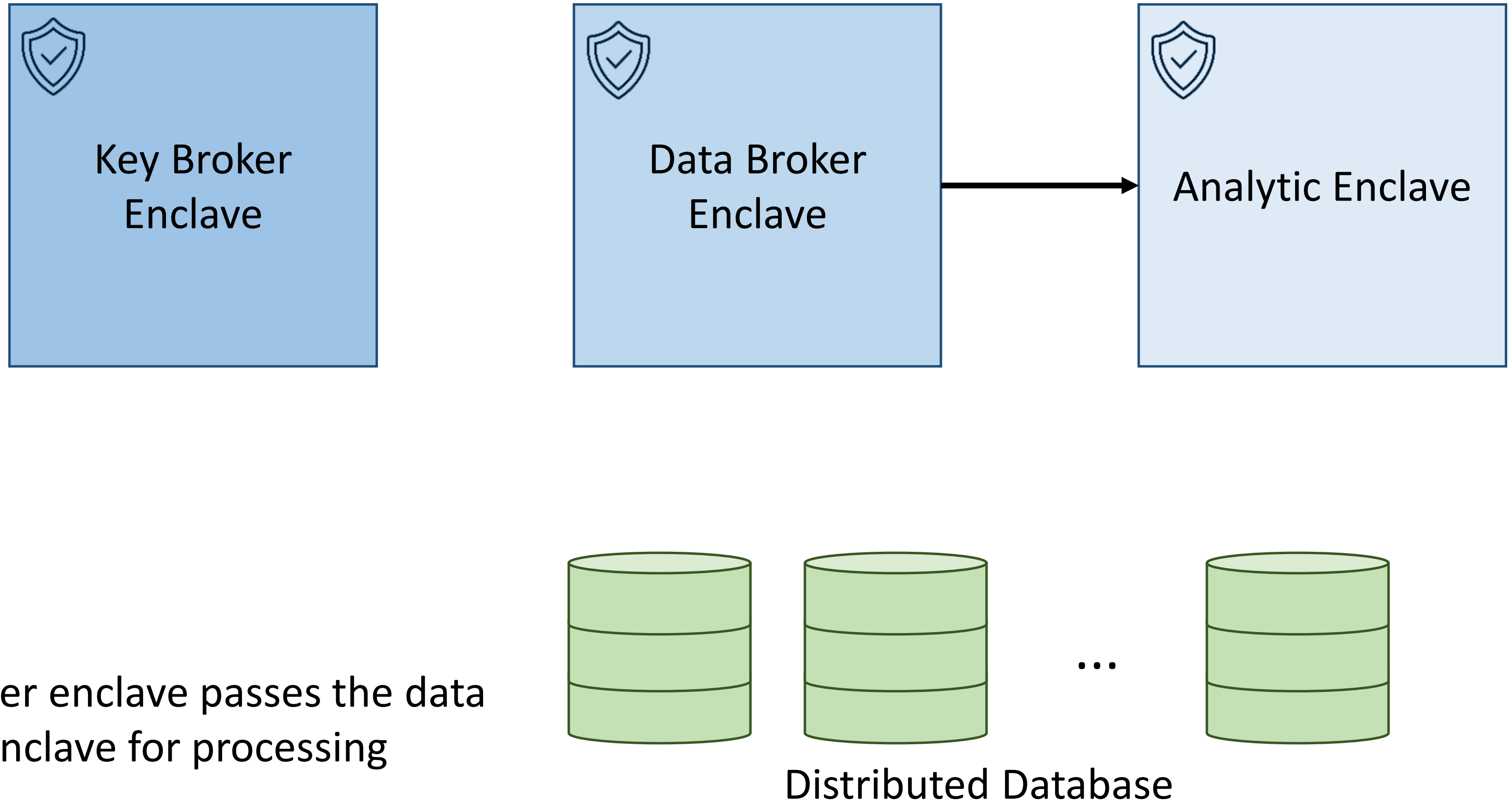
c. The key broker unwraps the file encryption key and passes it to the data broker enclave

# Database Queries



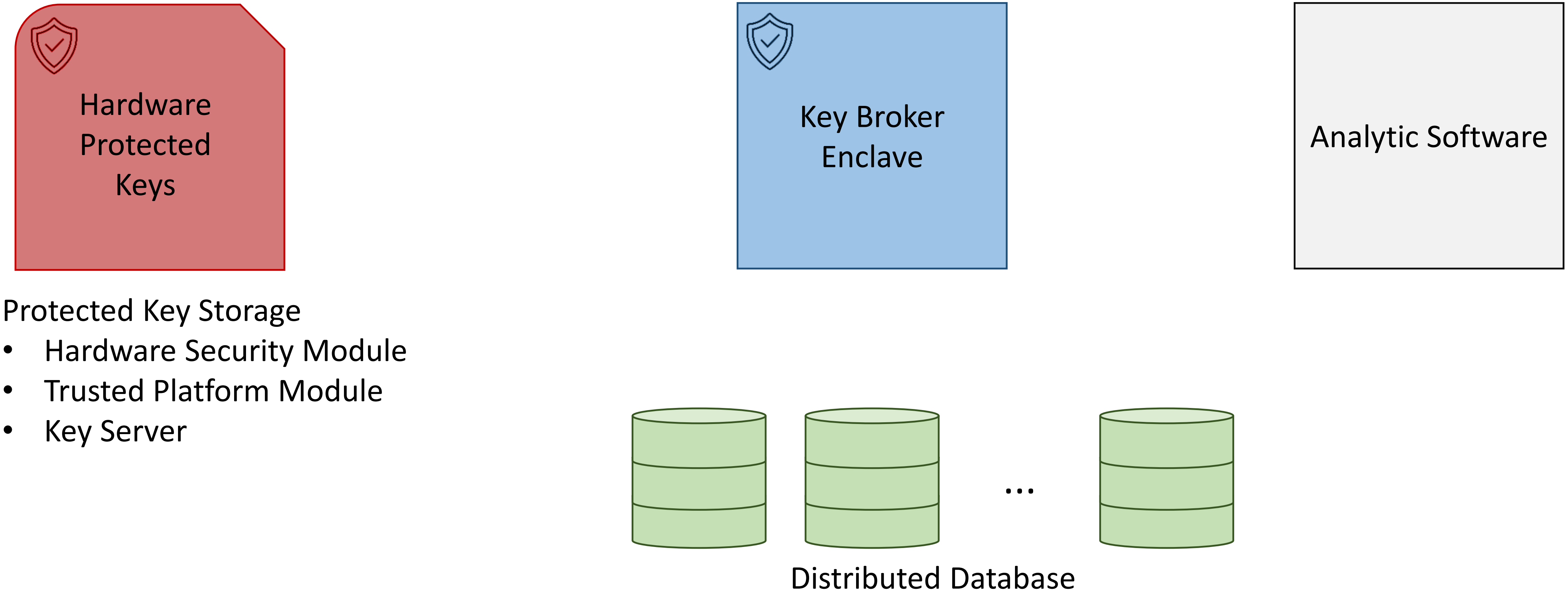
d. The data broker enclave decrypts the data

# Database Queries



e. The data broker enclave passes the data to the analytic enclave for processing

# Protecting Keys-in-Use



Hardware Protected Keys

Protected Key Storage

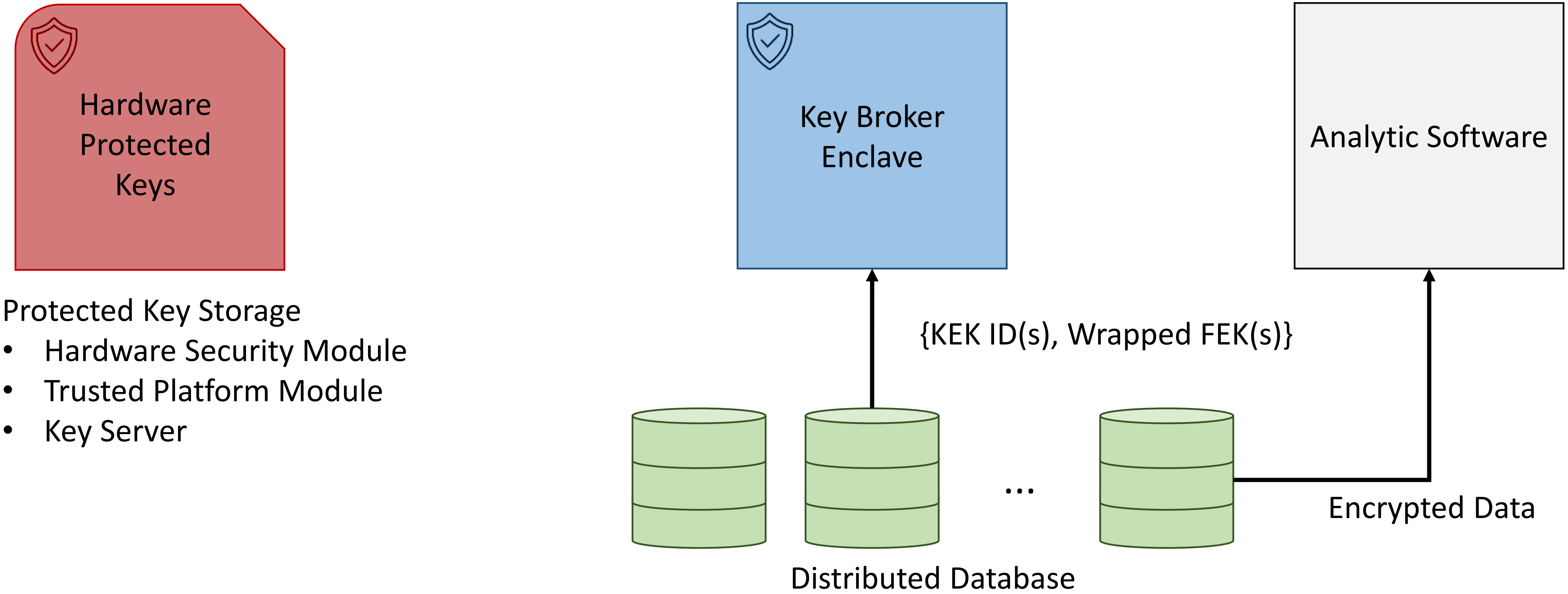
- Hardware Security Module
- Trusted Platform Module
- Key Server

Key Broker Enclave

Analytic Software

Distributed Database

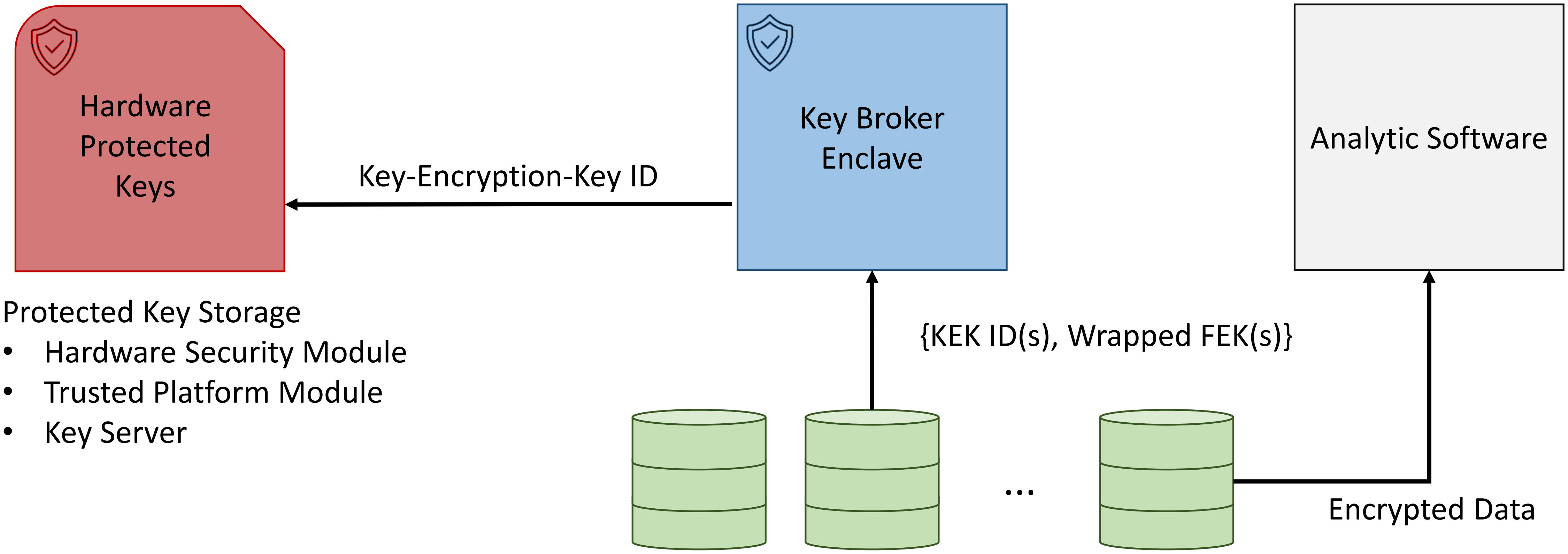
# Protecting Keys-in-Use



Protected Key Storage

- Hardware Security Module
- Trusted Platform Module
- Key Server

# Protecting Keys-in-Use

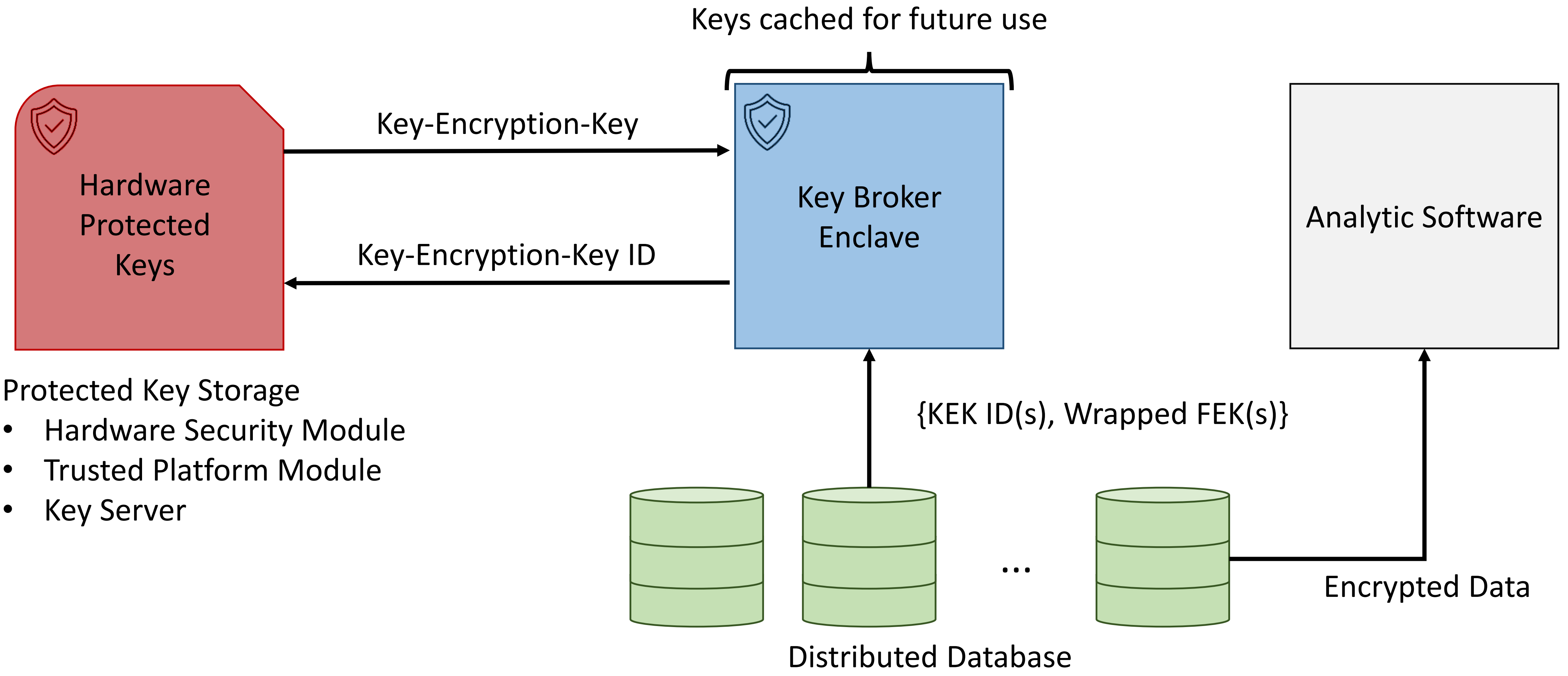


Distributed Database

Protected Key Storage

- Hardware Security Module
- Trusted Platform Module
- Key Server

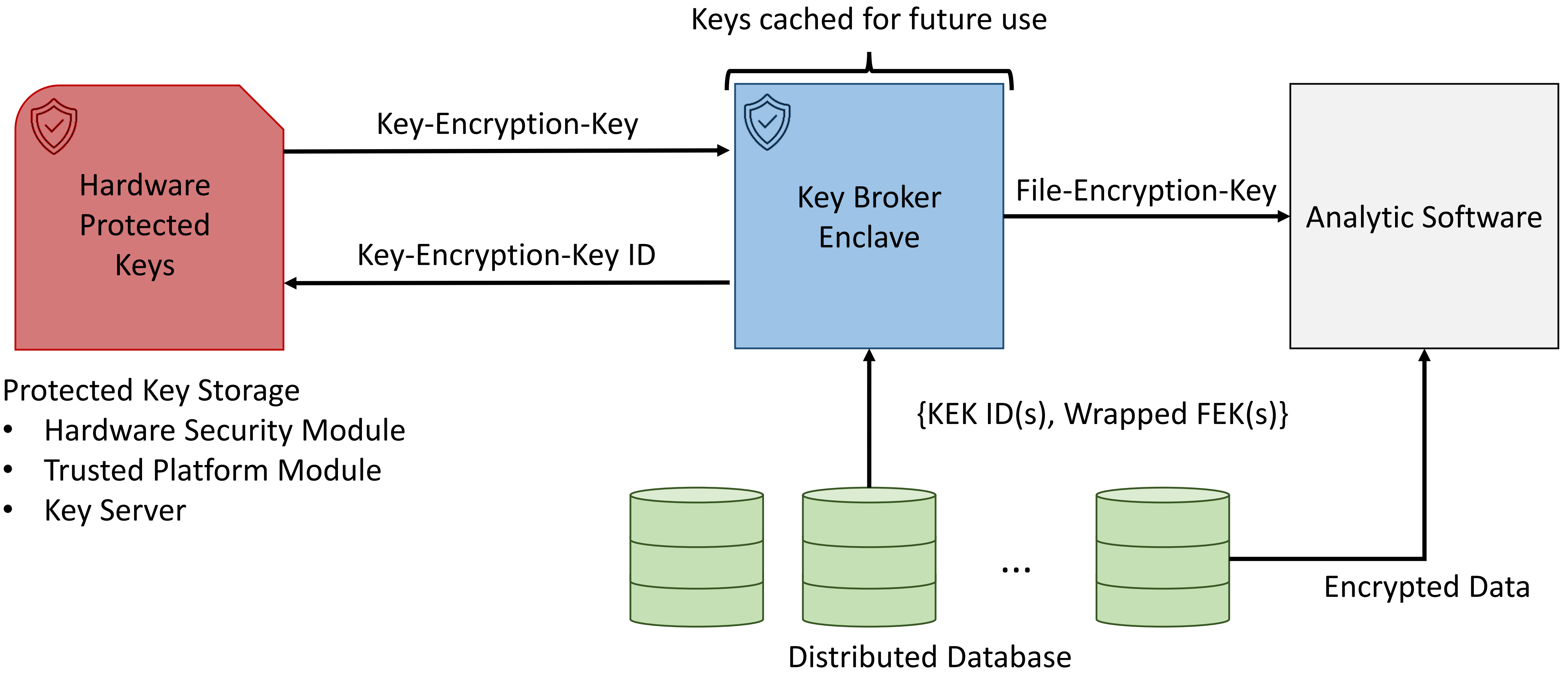
# Protecting Keys-in-Use



Protected Key Storage

- Hardware Security Module
- Trusted Platform Module
- Key Server

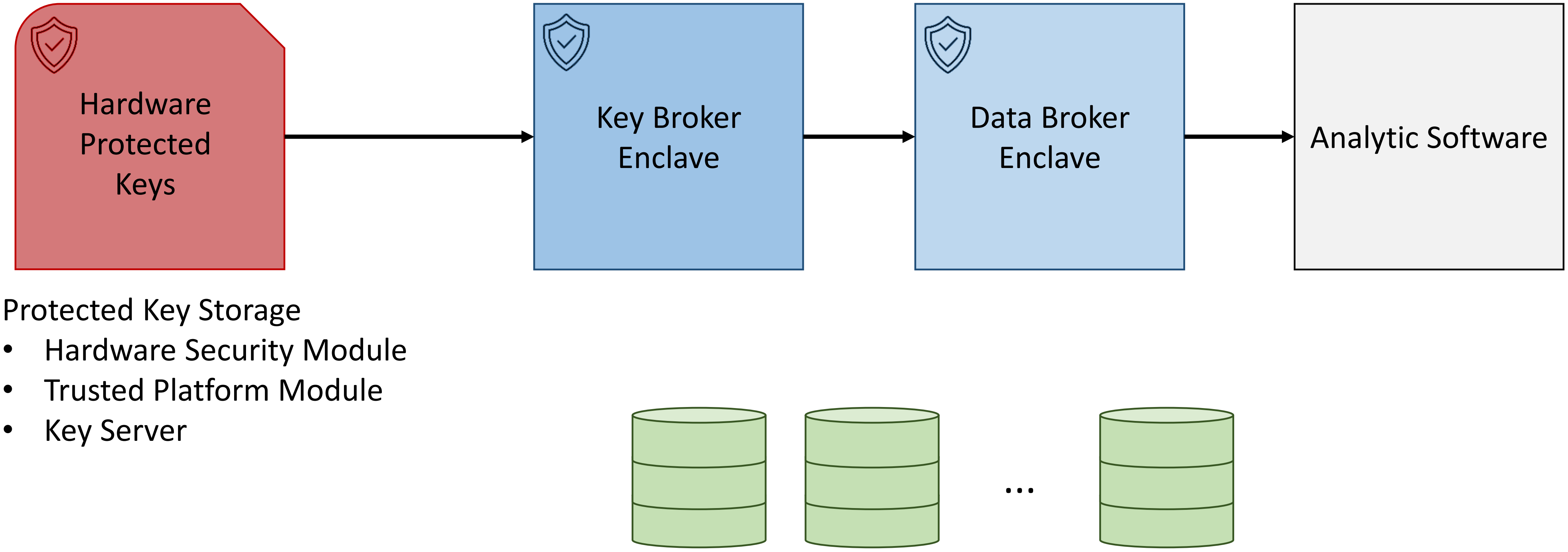
# Protecting Keys-in-Use



Protected Key Storage

- Hardware Security Module
- Trusted Platform Module
- Key Server

# Protecting Keys-in-Use (cont'd)

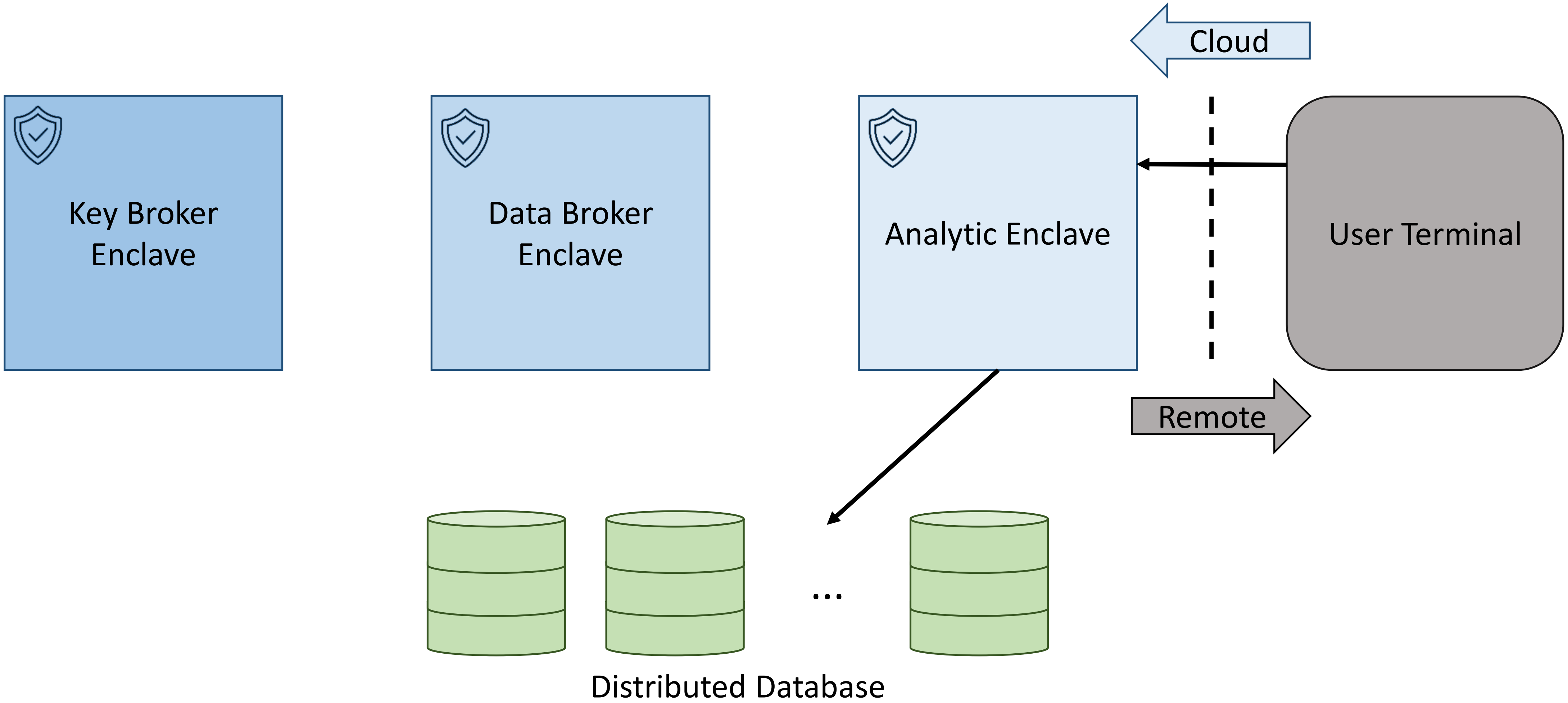


Protected Key Storage

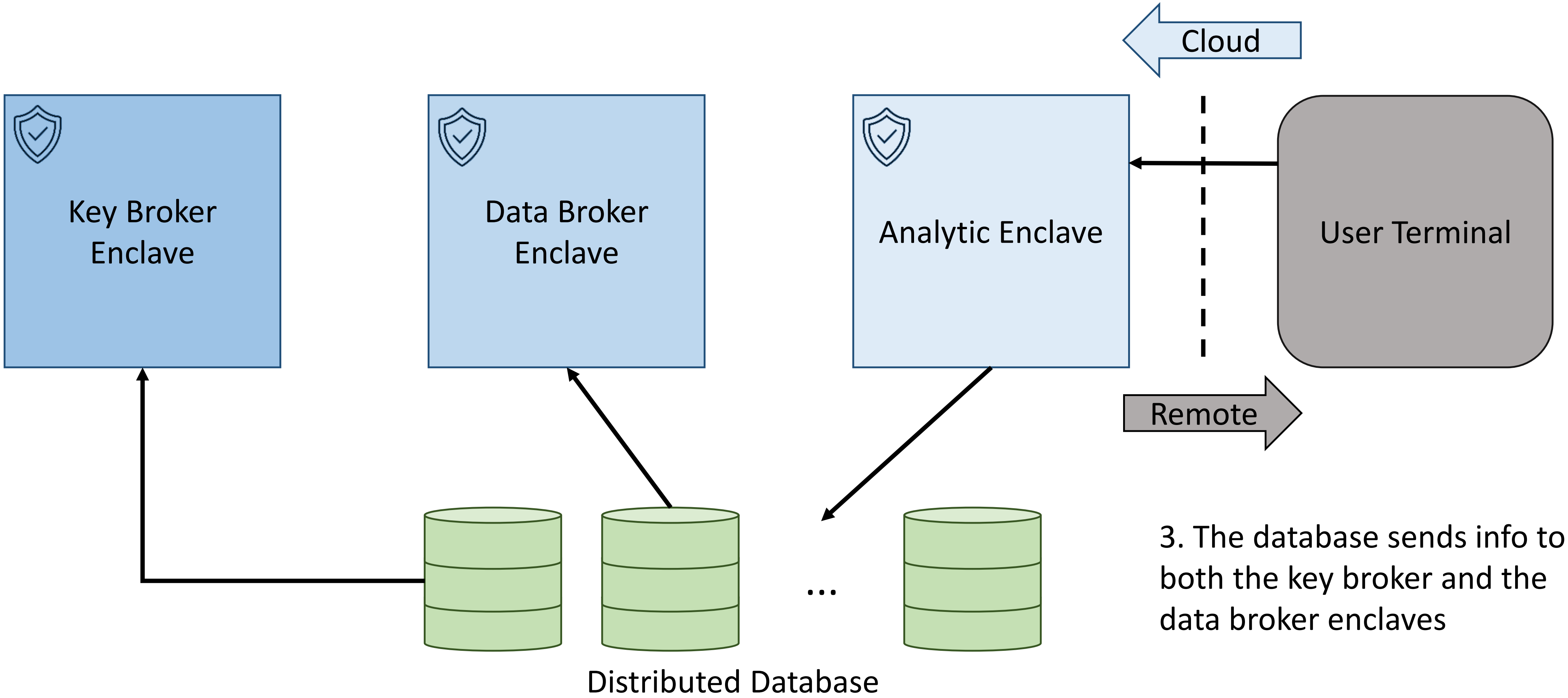
- Hardware Security Module
- Trusted Platform Module
- Key Server

Distributed Database

# Bird's-eye View

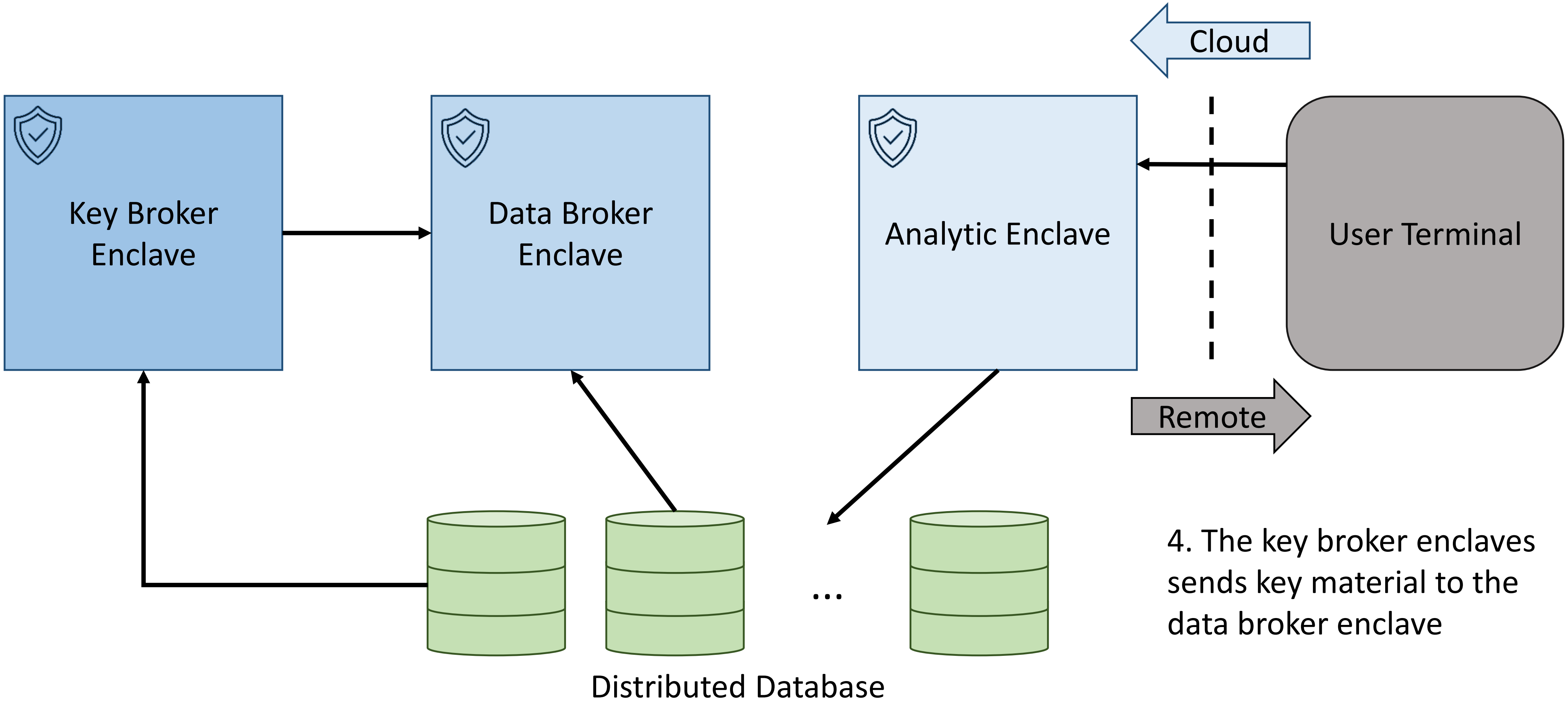


# Bird's-eye View



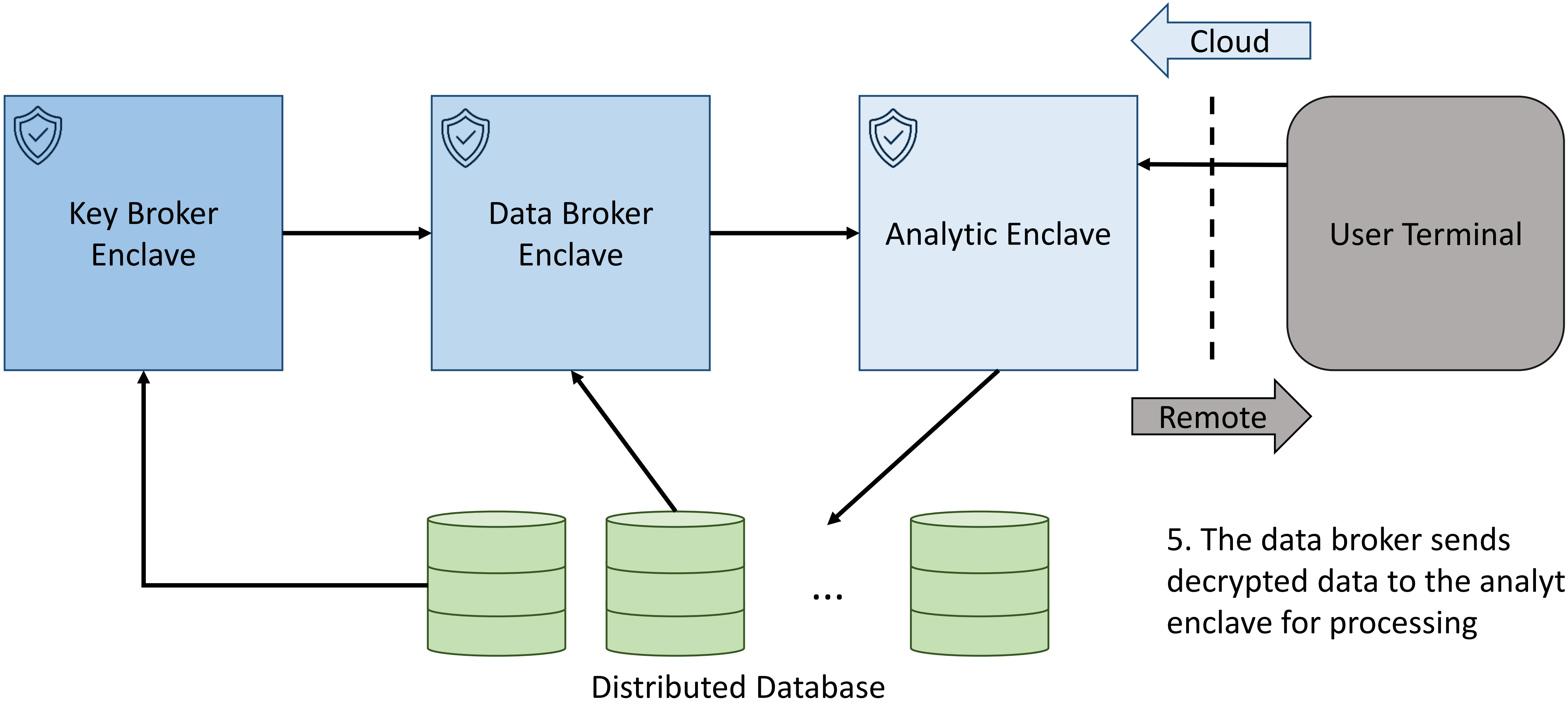
3. The database sends info to both the key broker and the data broker enclaves

# Bird's-eye View



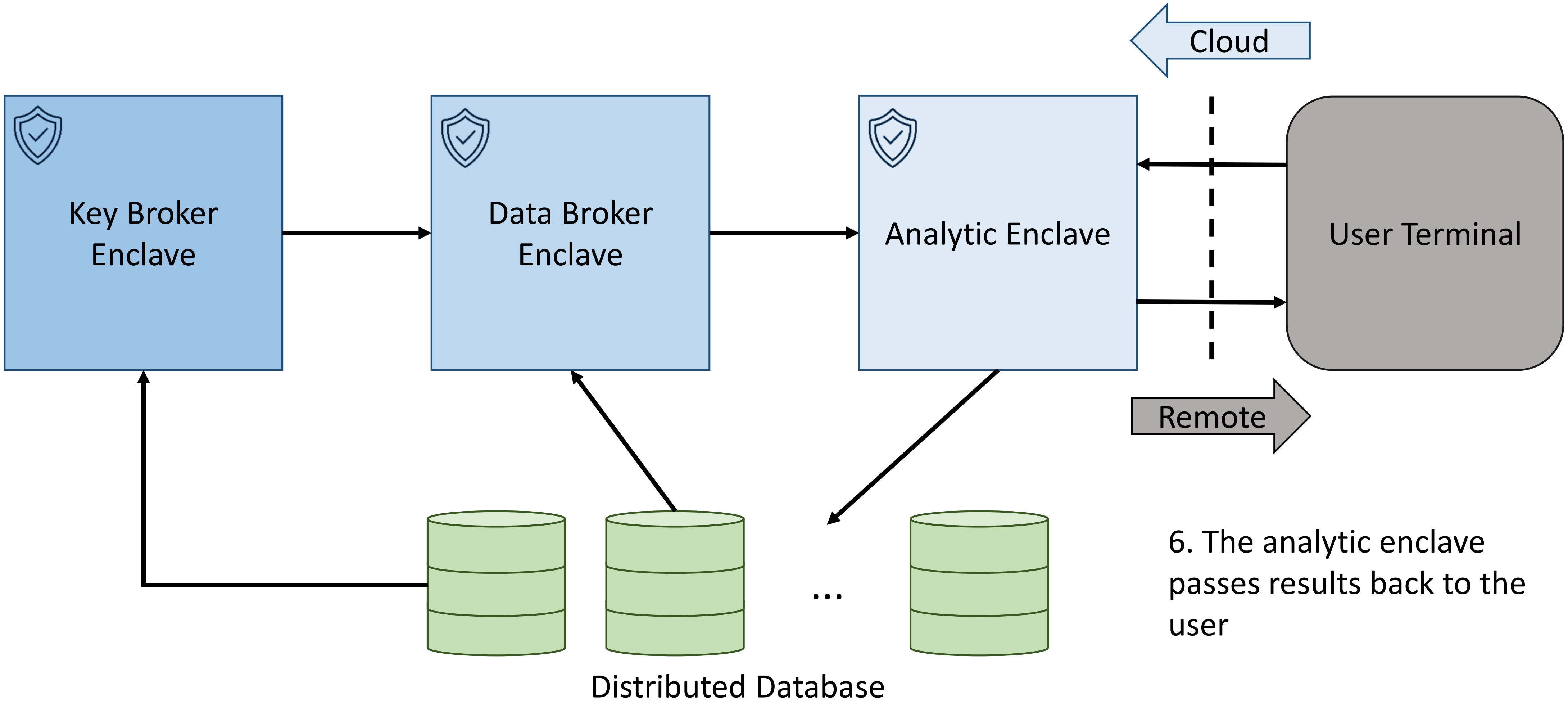
4. The key broker enclaves sends key material to the data broker enclave

# Bird's-eye View



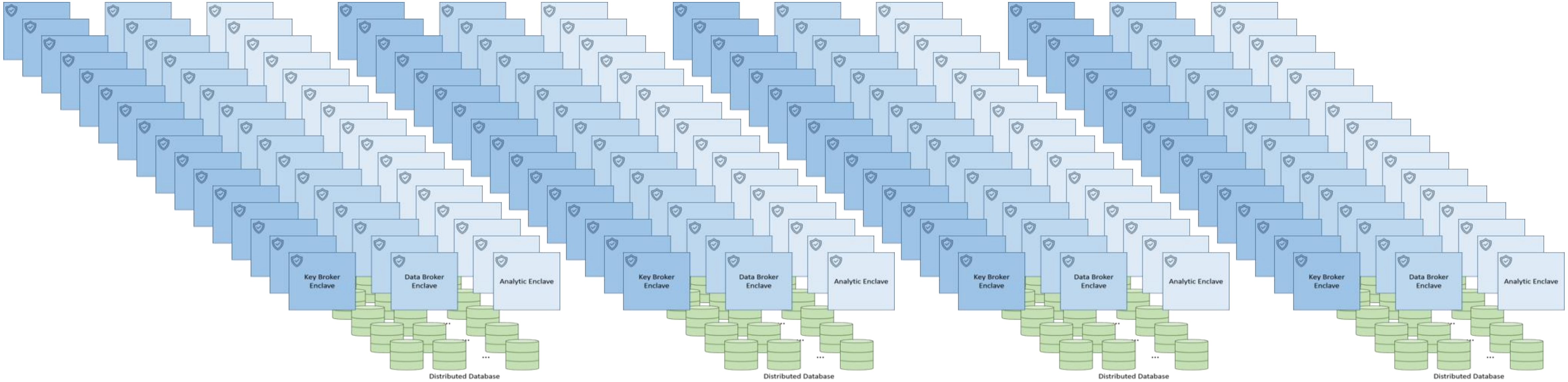
5. The data broker sends decrypted data to the analytic enclave for processing

# Bird's-eye View



6. The analytic enclave passes results back to the user

# Many Nodes within a Cloud



# Key Takeaways

- Trusted hardware has the potential to protect data-in-use within cloud computing
- Security is about making an adversary's job harder
- Protecting keys-in-use requires less powerful enclaves and provides meaningful protections
- Protecting keys-in-use paves the way for data-in-use
    - Both in capabilities and knowledge
    - Becomes more practical as secure enclaves become more powerful

# Current Work

- Key broker prototype:
  - Open sourced TPM 2.0 tool kmyth
  - Open sourced pelz key manager
  - Performance improvement over raw Java
  - Working towards TEE (SGX) prototype

- https://github.com/NationalSecurityAgency/kmyth
- https://github.com/NationalSecurityAgency/pelz

# References

[1] BMC Blogs. ”Gartner 2020 Magic Quadrant for Cloud Infrastructure and Platform Services - BMC Blogs.” https://www.bmc.com/blogs/gartner-magic-quadrant-cips-cloud-infrastructure-platform-services/ (accessed Jan. 29, 2021).

[2] Y. Nawaz, Keynote Presentation, Topic: "Building Secure Financial Platforms in the Cloud." ACM Cloud Computing Security Workshop, London, Nov. 11, 2019.

[3] M. Russinovich, Confidential Computing. (Aug. 27, 2018). Accessed: Sep. 2, 2020. [Online Video]. Available: https://youtube.com/watch?v=SUS3Zzko3eM

[4] R. Chandramouli, M. Iorga, and S. Chokhani, "Cryptographic Key Management Issues & Challenges in Cloud Services," National Institute of Standards and Technology, Gaithersburg, MD, USA, NISTIR 7956, Sep. 2013. [Online]. Available: https://csrc.nist.gov/publications/detail/nistir/7956/final

[5] Microsoft. “Confidential computing on a healthcare platform.“ Microsoft. https://docs.microsoft.com/en-us/azure/architecture/example-scenario/confidential/healthcare-inference#architecture (accessed Feb. 3, 2021).