!Oxymoron

Computing on Encrypted Data

Srinivasan Narayanamurthy (Srini)

May 26th, 2017

Storage Developers Conference India, Bangalore

NetApp

About

- Me
  - Engineer at the Advanced Technology Group since 2011.
  - Spent a decade working on security before joining NetApp.
    - Mostly on a PhD & several years at RSA Security.
  - @NetApp: Data Security & Privacy, Erasure codes, Distributed Storage Systems.
- My Involvement in SNIA
  - SDC talks (at India & US) in 2016 on Erasure Codes.
  - Member of SNIA Security TWG.
    - Tries hard to stay awake until 2:30AM (IST) to attend weekly meetings! ☺
  - Early version of this talk at Data Storage Security Summit 2016.
- Talk: Searchable Encryption
  - Non-mathematical; non-algorithmic.

NetApp

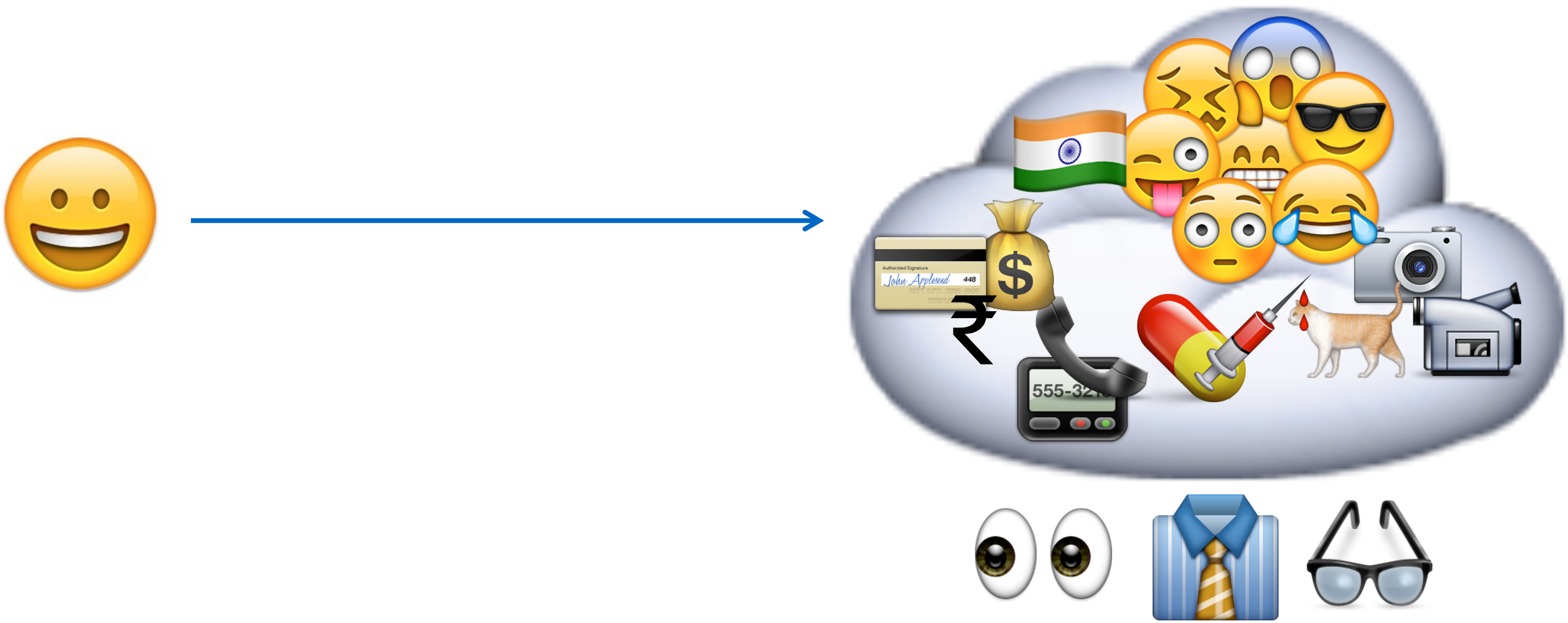
Oxymoron

[ok-si-mawr-on, -mohr-]

a figure of speech by which a locution produces an incongruous, seemingly self-contradictory effect, as in “cruel kindness” or “to rush slowly” OR “computing/ searching on encrypted data.”

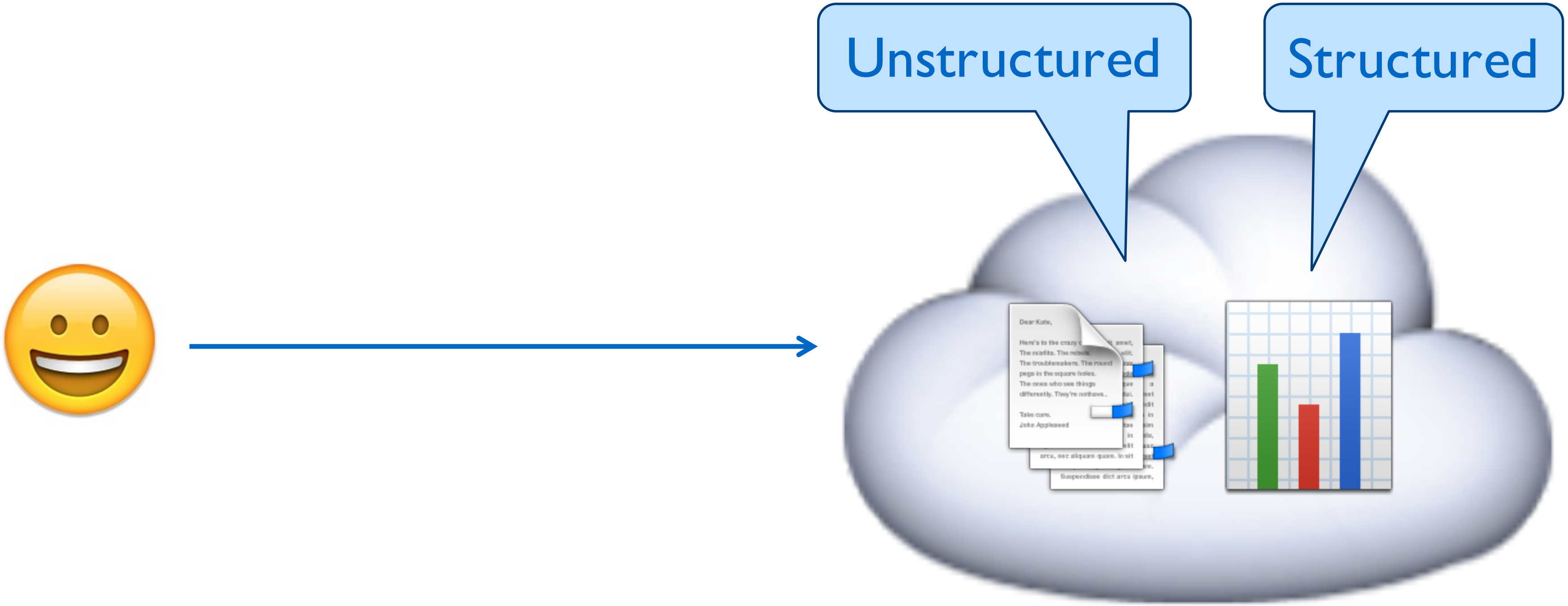
NetApp

The Problem

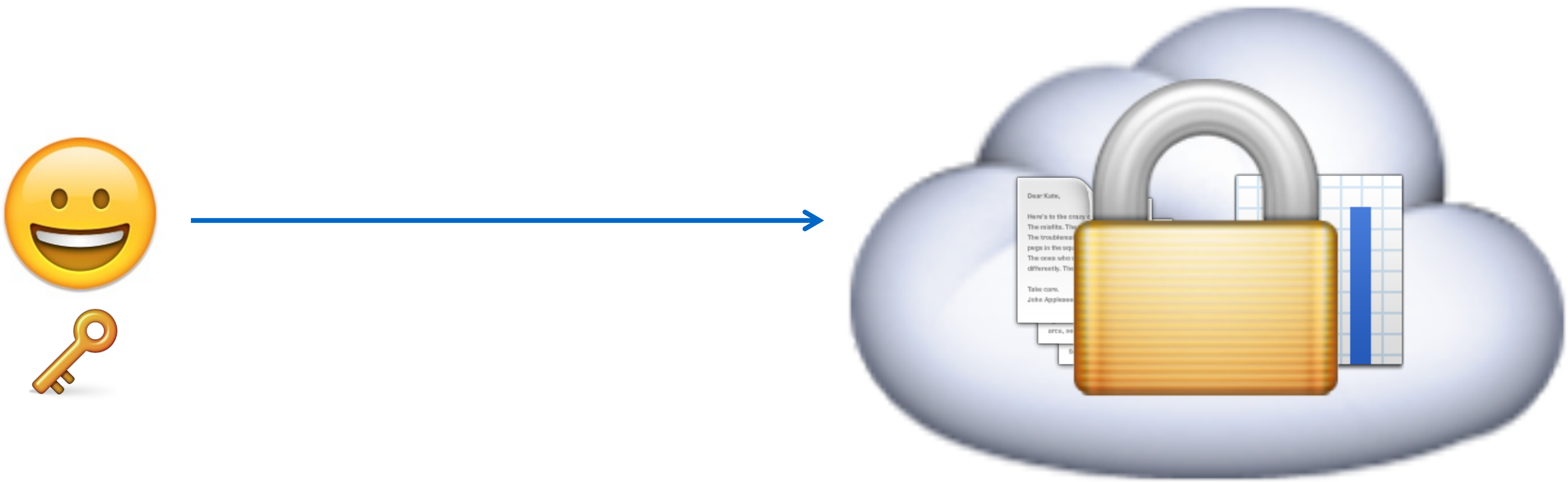


Semi-trusted (Honest-but-curious) server

Classification

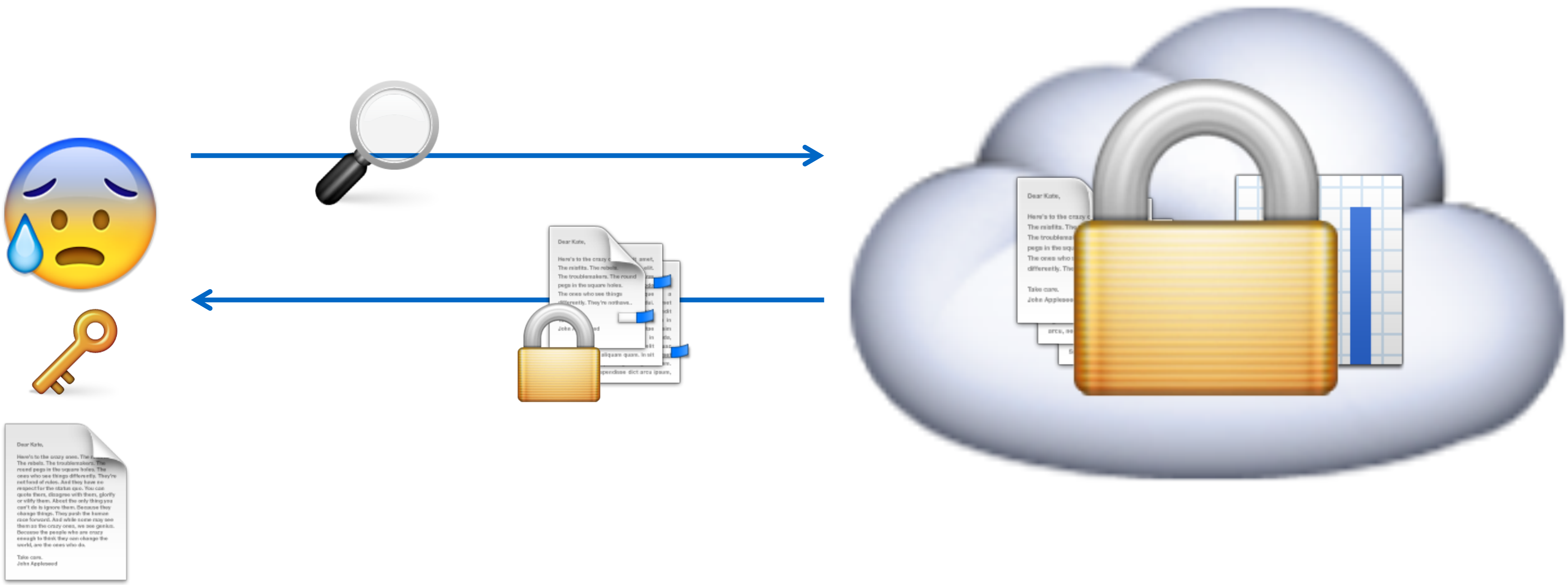


Solution

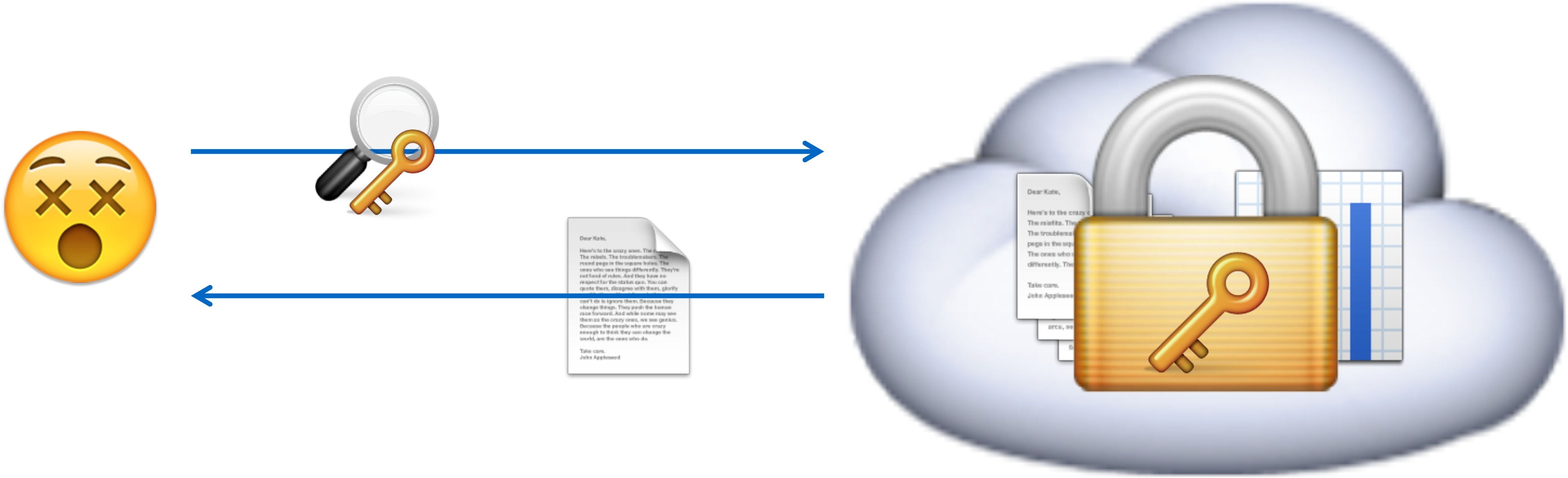


Encrypt !

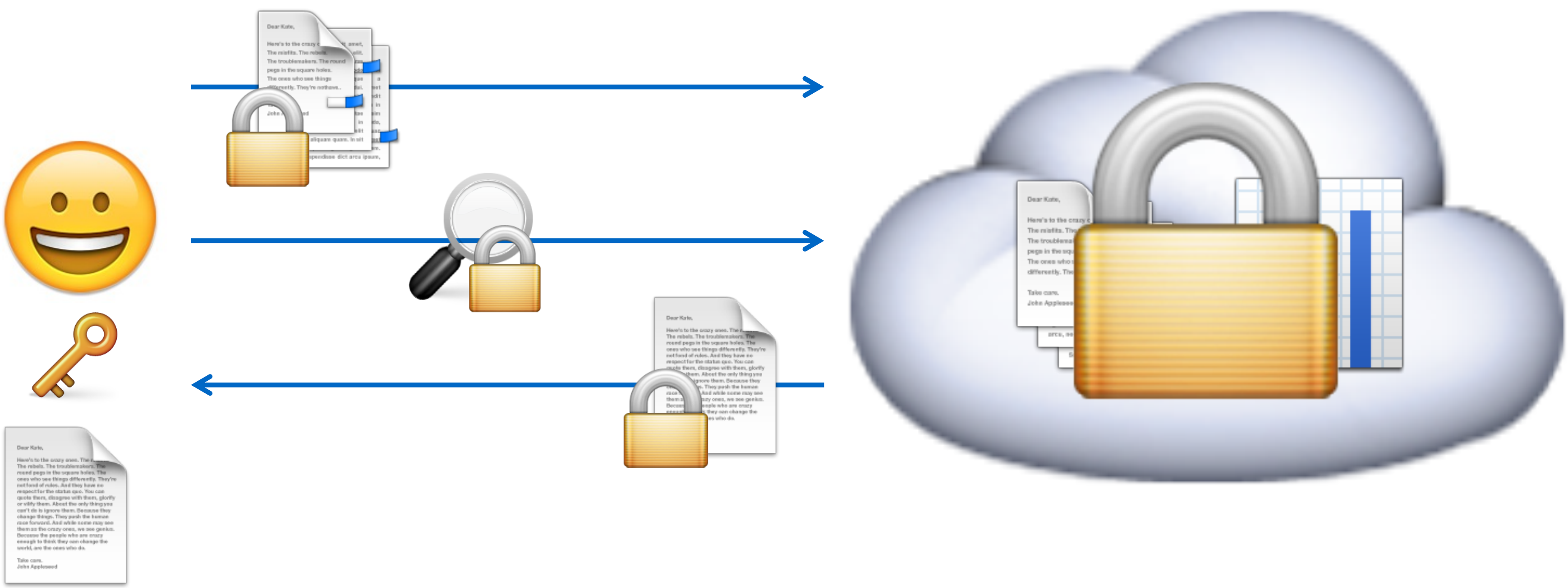
But …



Or …



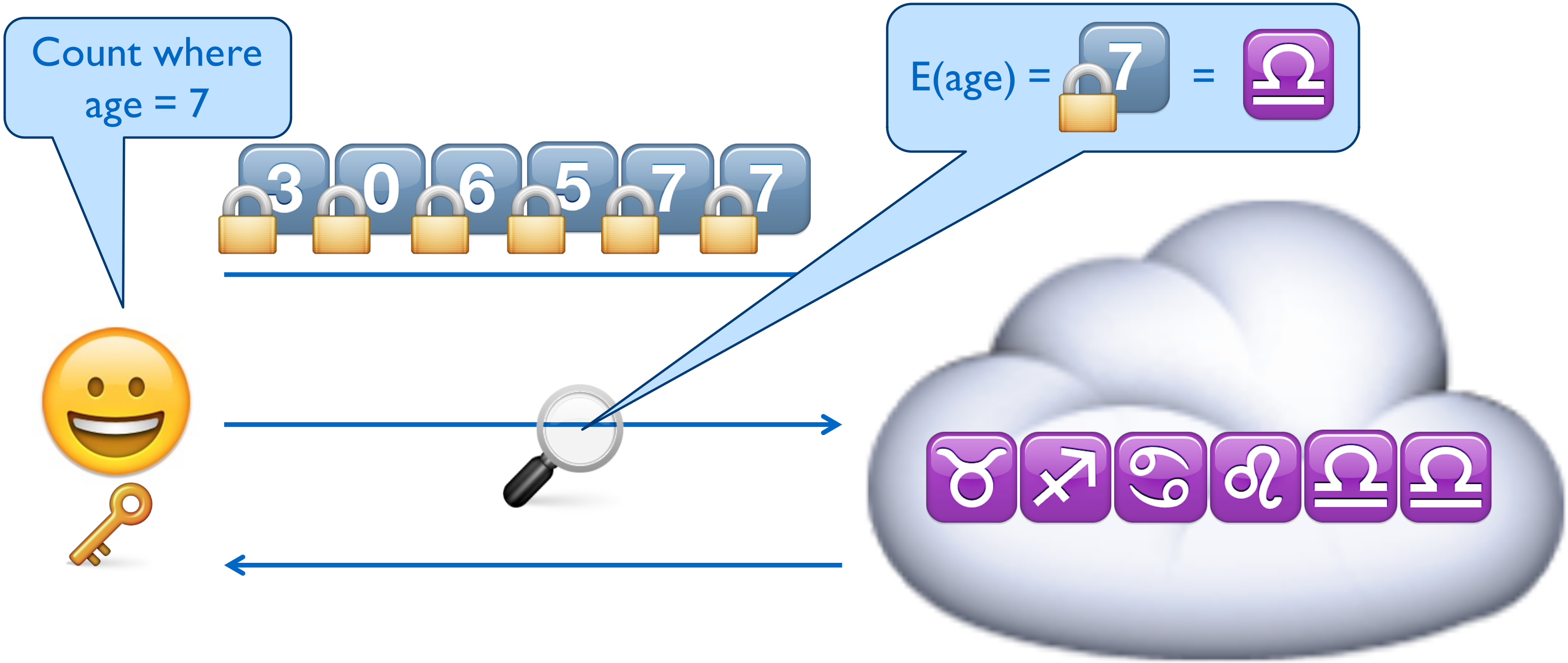
Searchable Encryption



Encrypted Data-at-rest & Data-in-motion

NetApp

Deterministic Encryption



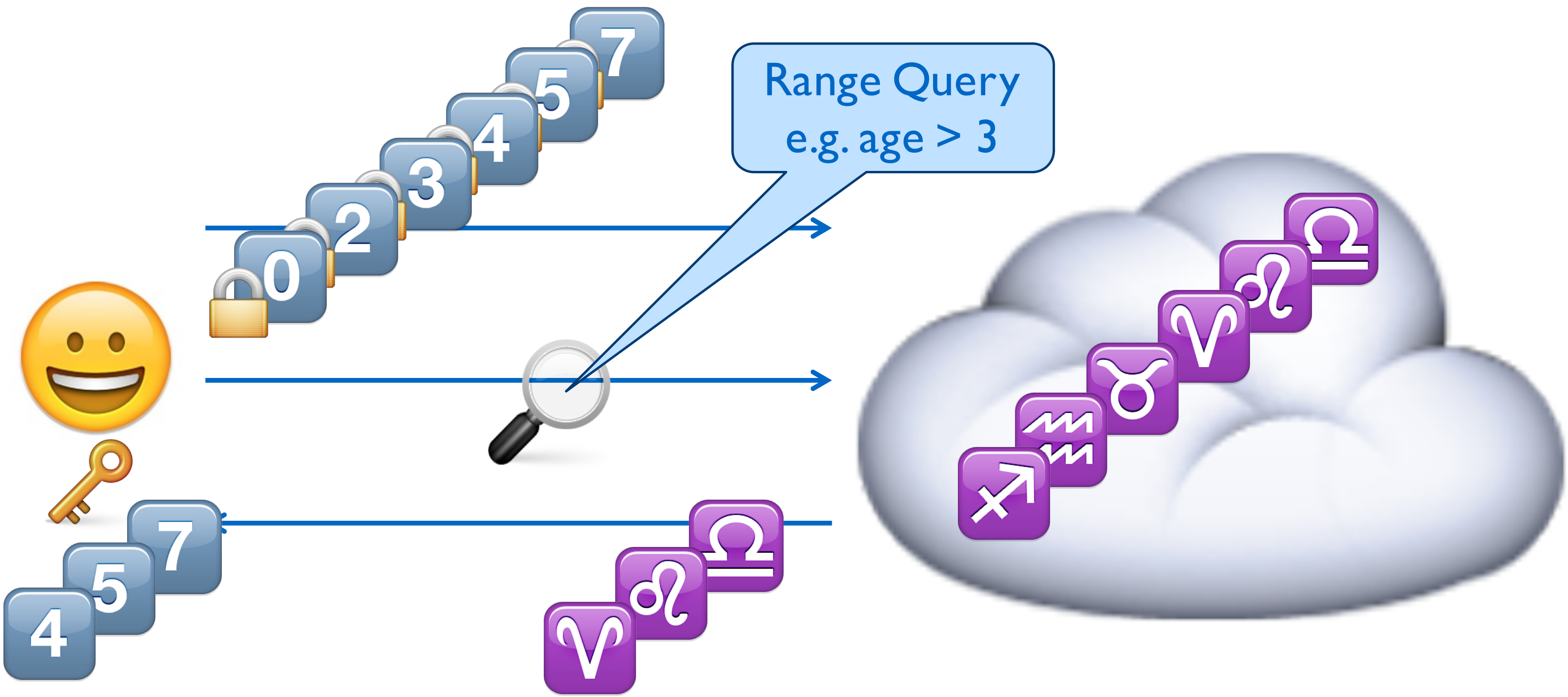
Example: AES – ECB mode

Application: Convergent Encryption for Deduplication

Brute-force / Dictionary attacks (IND-CPA)

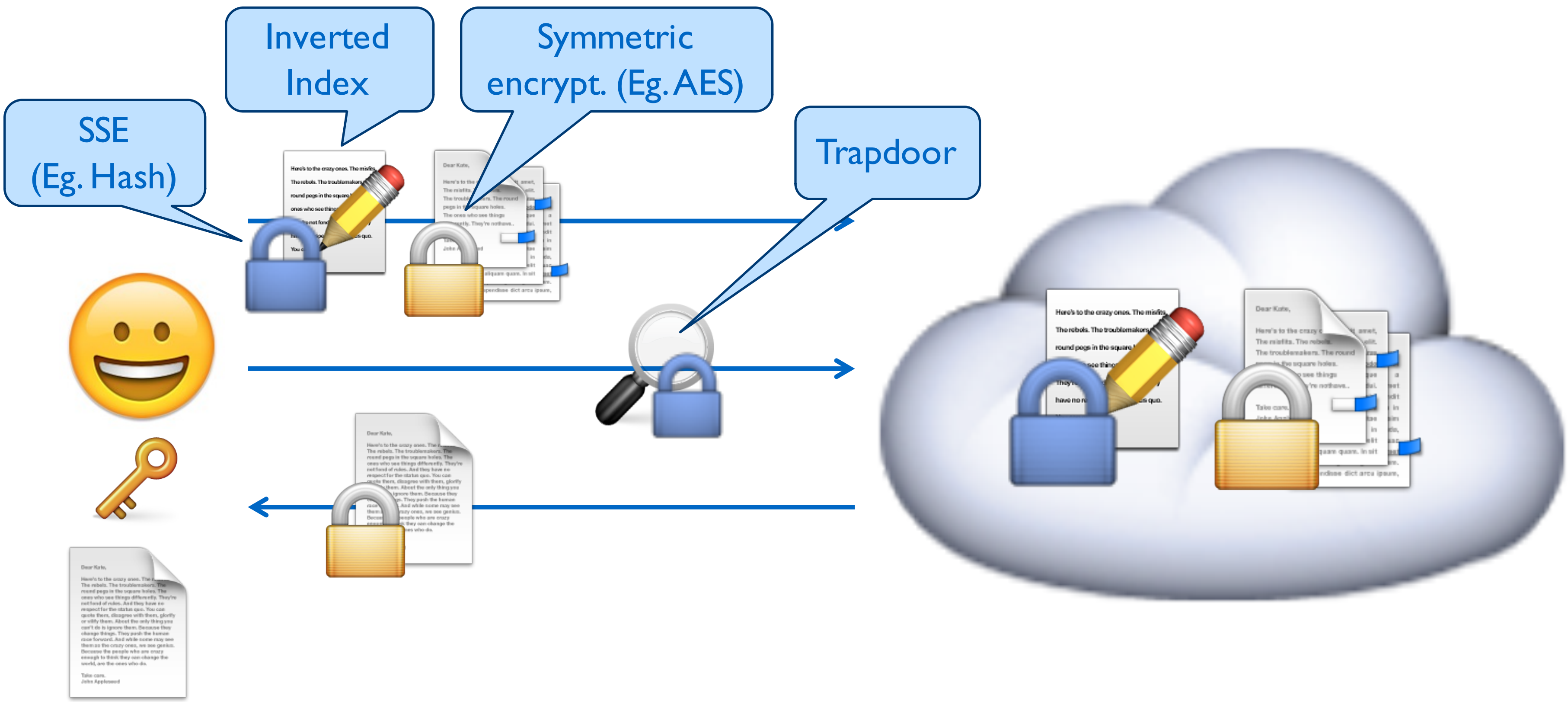
NetApp

Order-preserving Encryption (OPE)



Symmetric encryption over integers (AES – FFX)

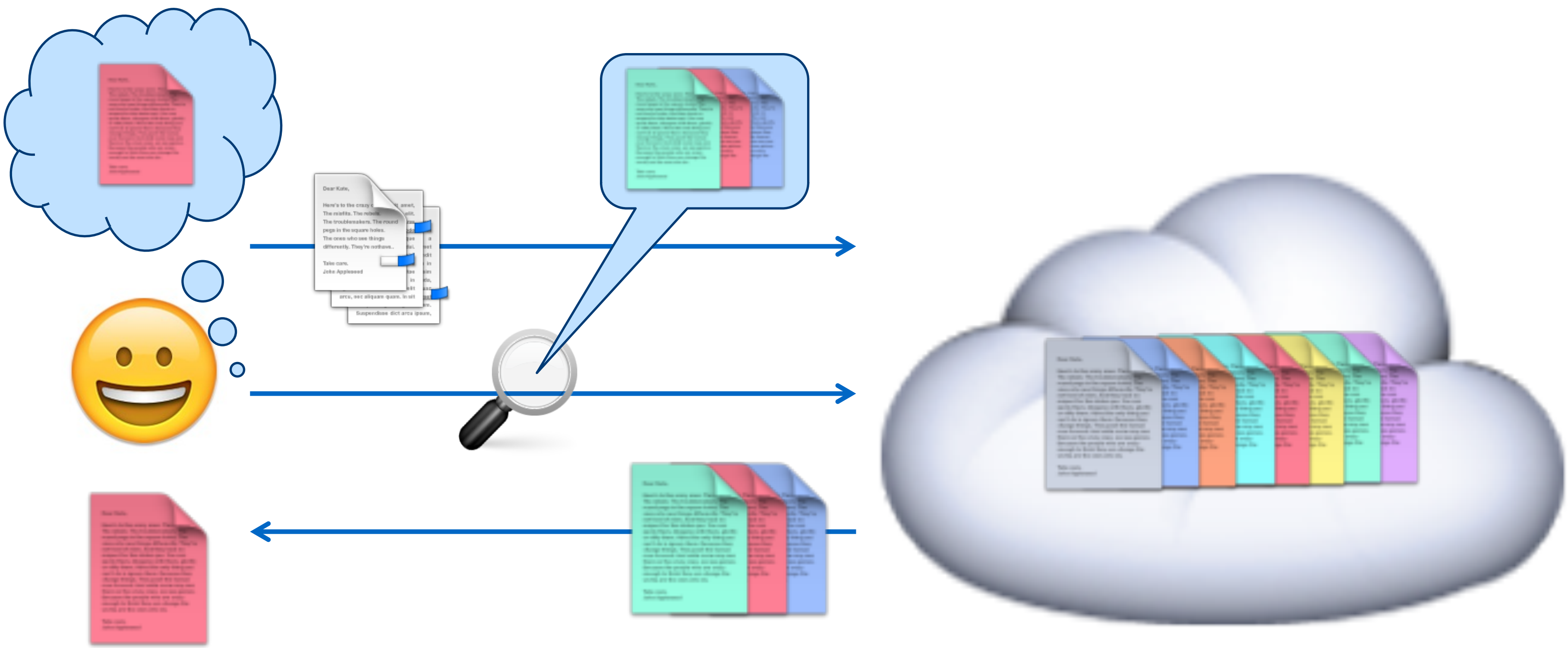
Searchable Symmetric Encryption (SSE)



Access pattern leakage!

NetApp

Oblivious RAM (oRAM)

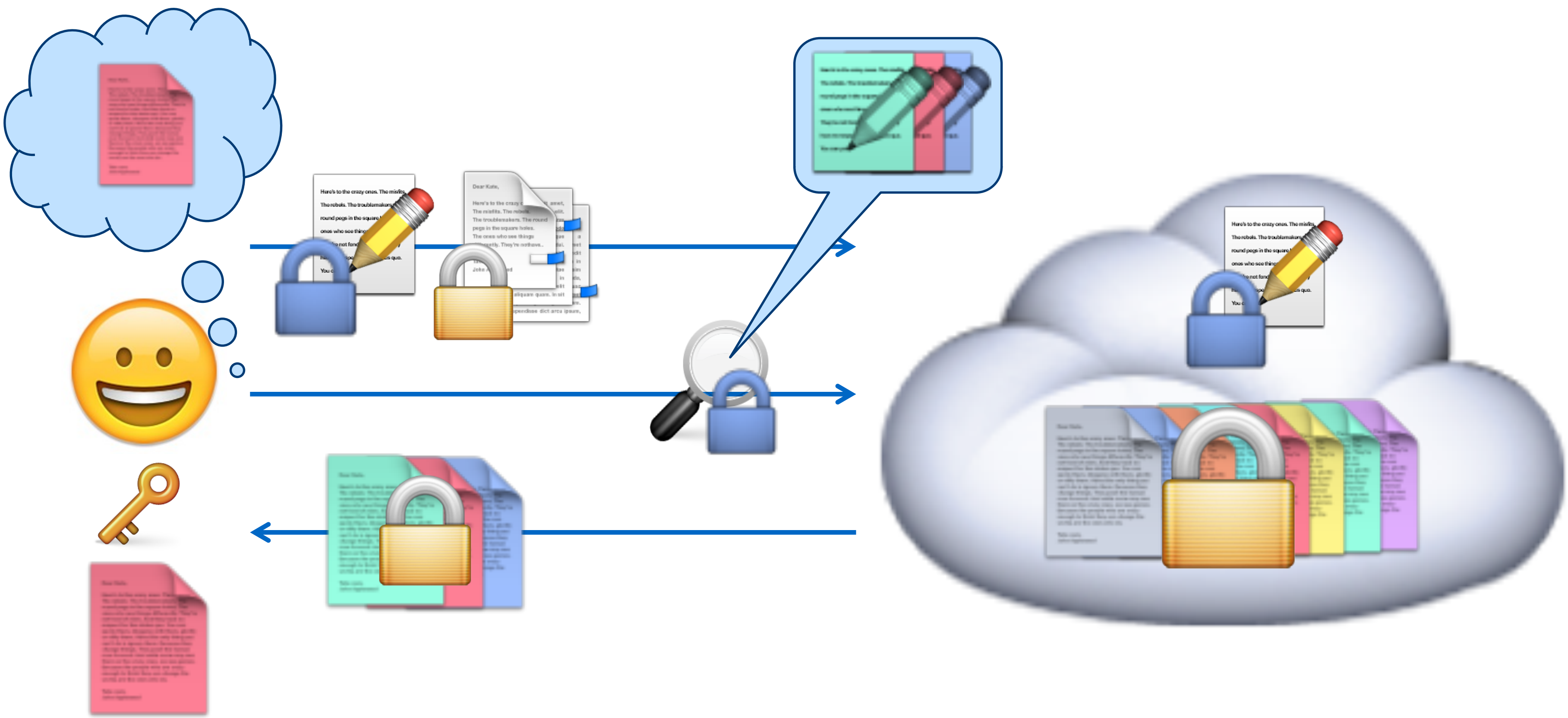


Hides all information, including access pattern

Many rounds of communication; Large storage cost

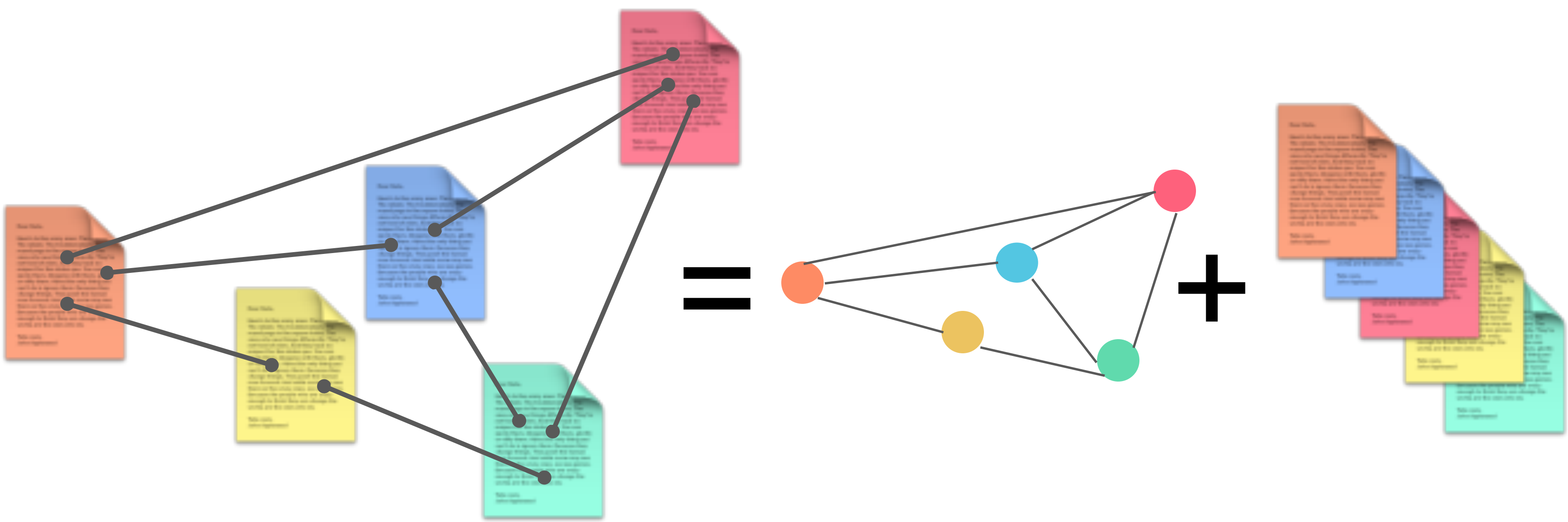
NetApp

SSE + oRAM



Hides all information, including access pattern;
Many rounds of communication, Large storage cost

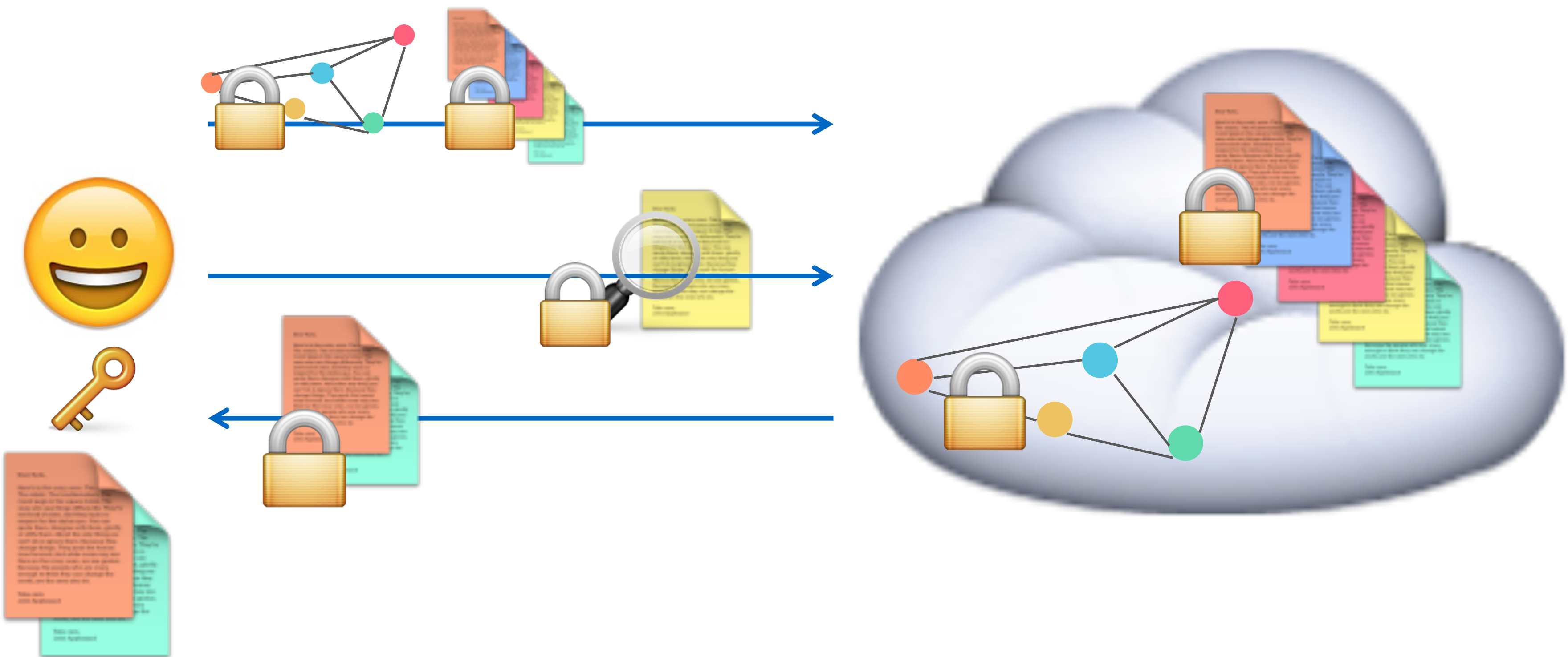
Structured Data



Social networks, Web crawlers, Maps, Network routing,
Communication (email headers, phone logs),
Research papers (citations)

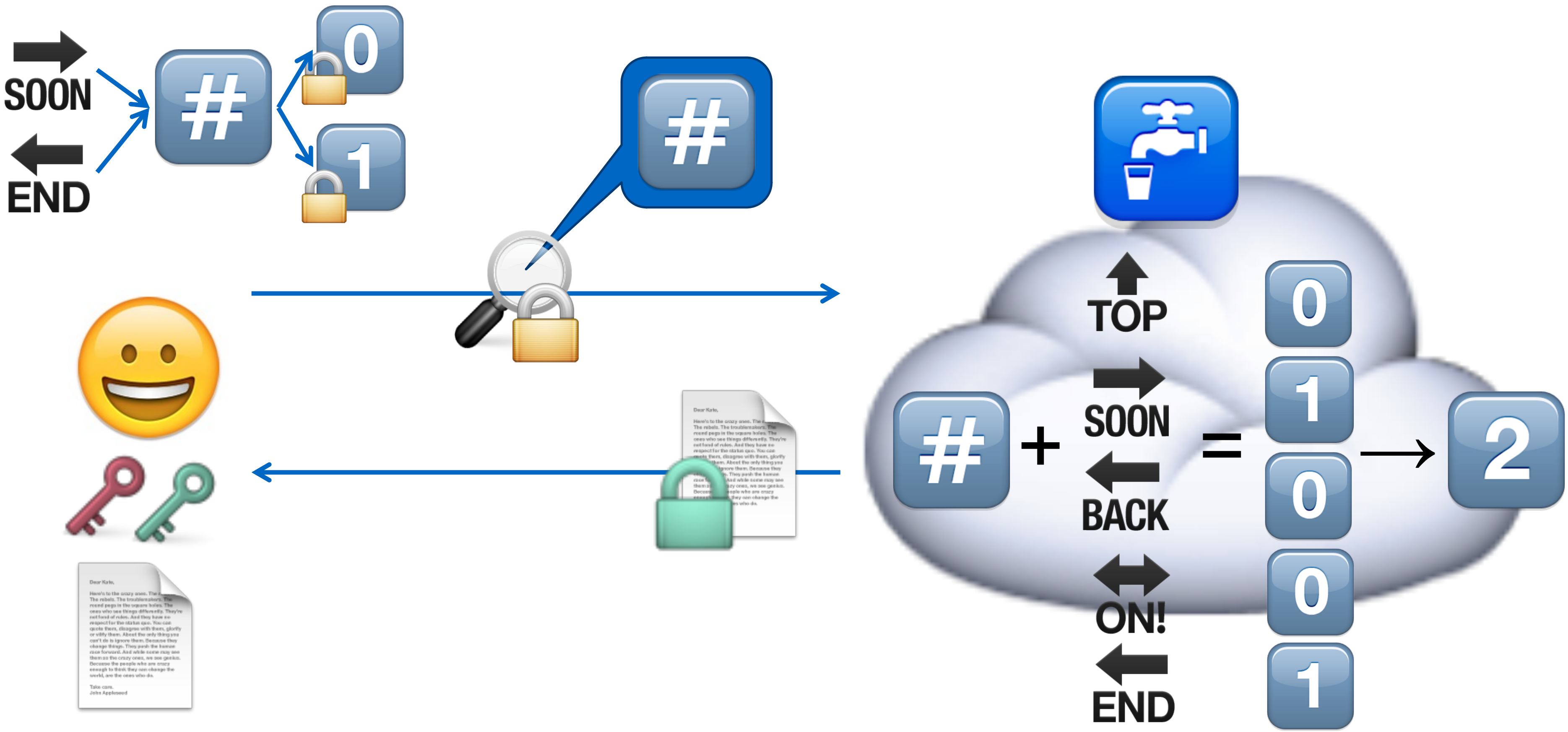
NetApp

Structured Encryption (STE)



NetApp

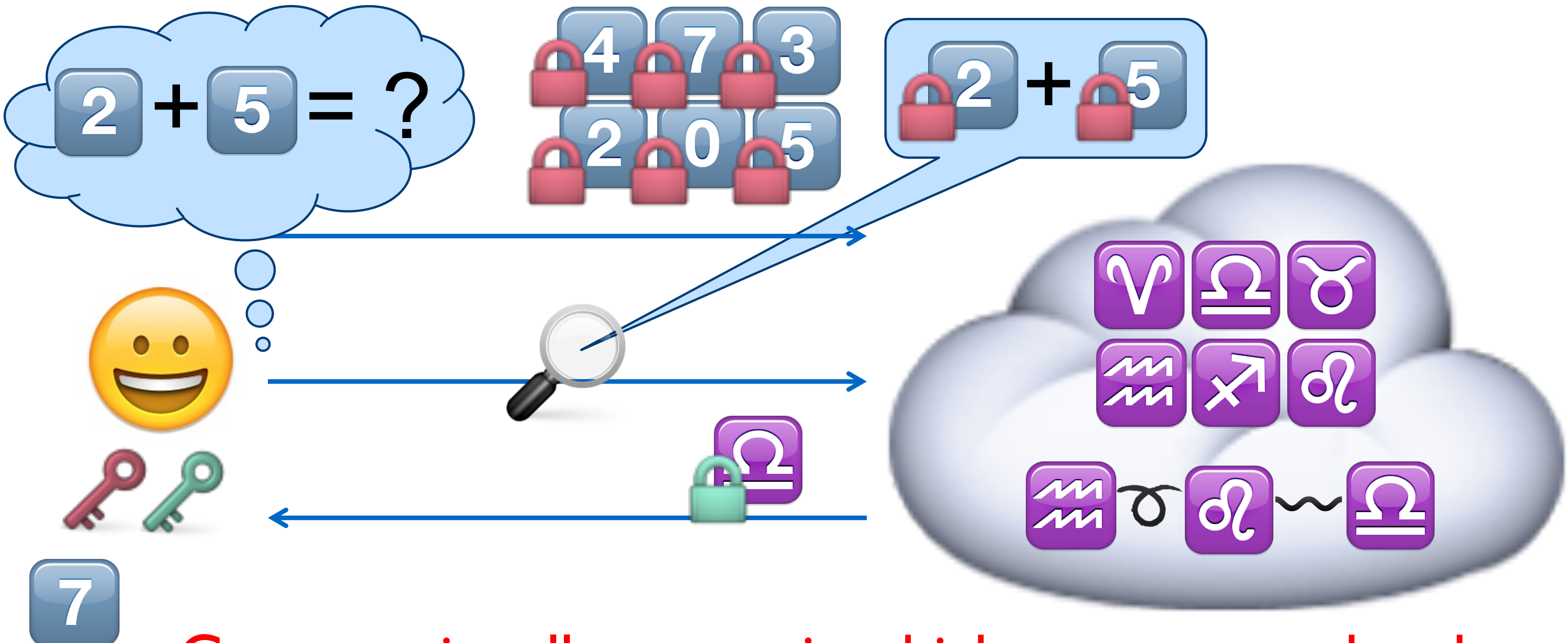
Private Stream Searching (PSS)



Partial (Additive) homomorphism!

NetApp

Fully Homomorphic Encryption (FHE)



Computationally expensive, high storage overhead

Search time is linear in the length of the dataset

Somewhat Homomorphic (SWHE):

Efficient; restricted number of additions and multiplications

NetApp

Other Encryption Schemes

- PKEET (Public Key Encryption with Equality Test)
  - Equality tests of plaintexts encrypted under different public keys
- PE (Predicate) & IPE (Inner Product)
  - Access-control & (originally) equality tests
  - IBE (Identity), AIBE (Anonymous IBE), HIBE (Hierarchical)
  - ABE (Attribute)
- HVE (Hidden Vector)
  - Wild card characters inside a key
  - Supports: conjunctive, subset, range queries, disjunctions, polynomials, inner products

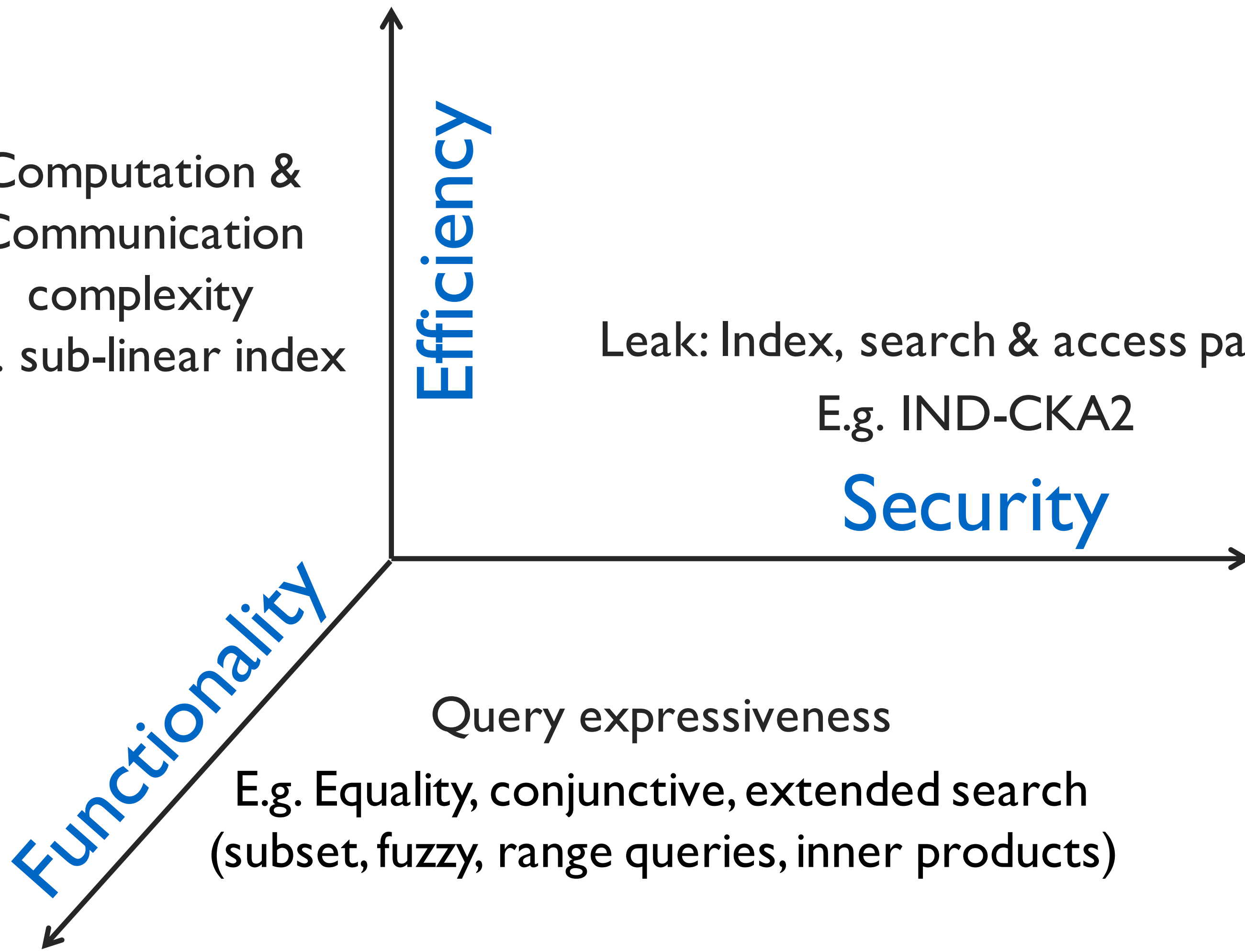
Summary

- Symmetric
  - Searchable Symmetric Encryption (SSE)
  - IND-CKA2 security
  - Efficient (sub-linear) SE schemes
- Asymmetric
  - Public key Encryption with Keyword search (PEKS)
  - Efficiency and security?
  - Lack of query expressiveness

NetApp

Tradeoffs

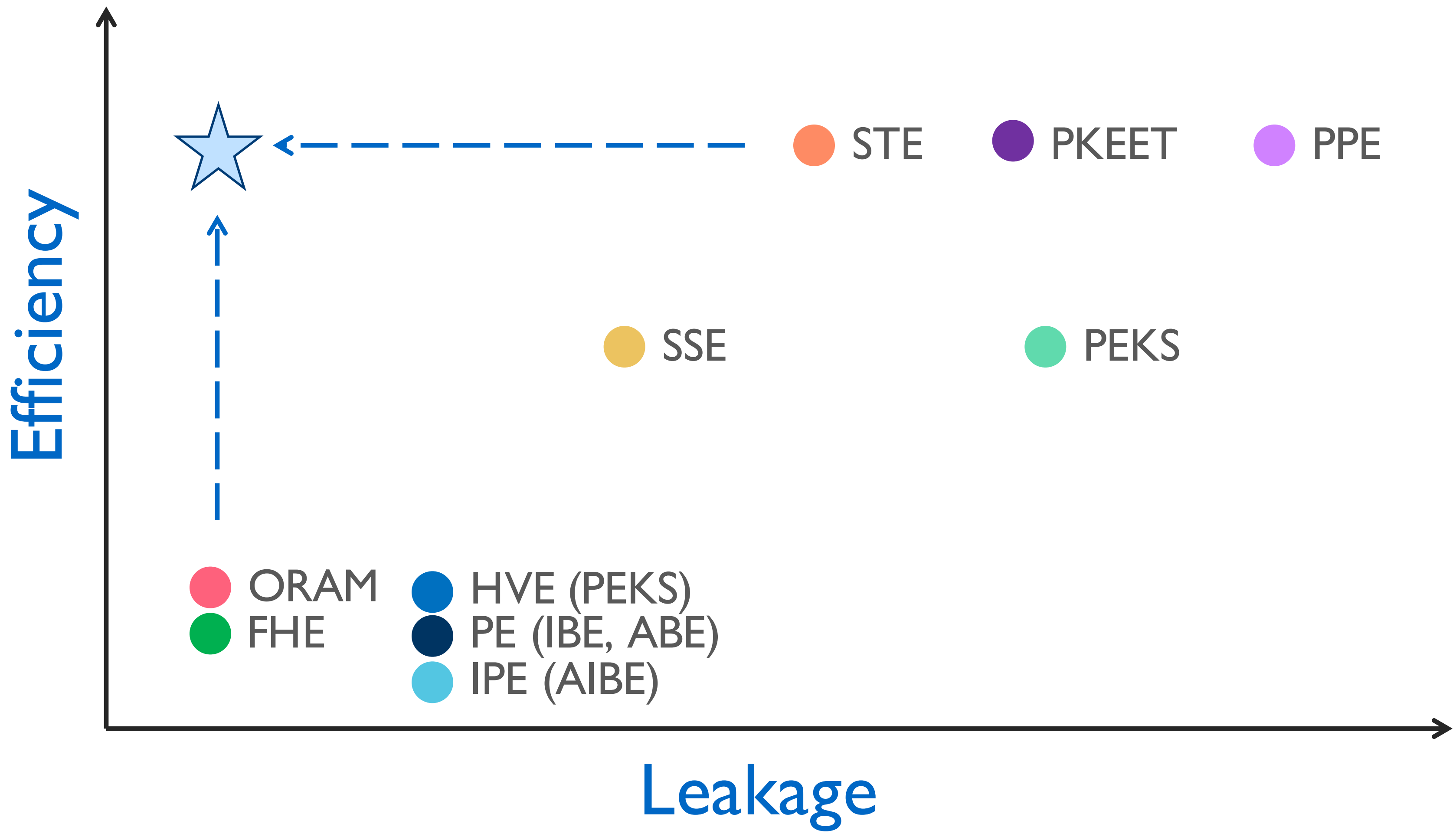
Computation &
Communication
complexity
E.g. sub-linear index



Leak: Index, search & access pattern
E.g. IND-CKA2

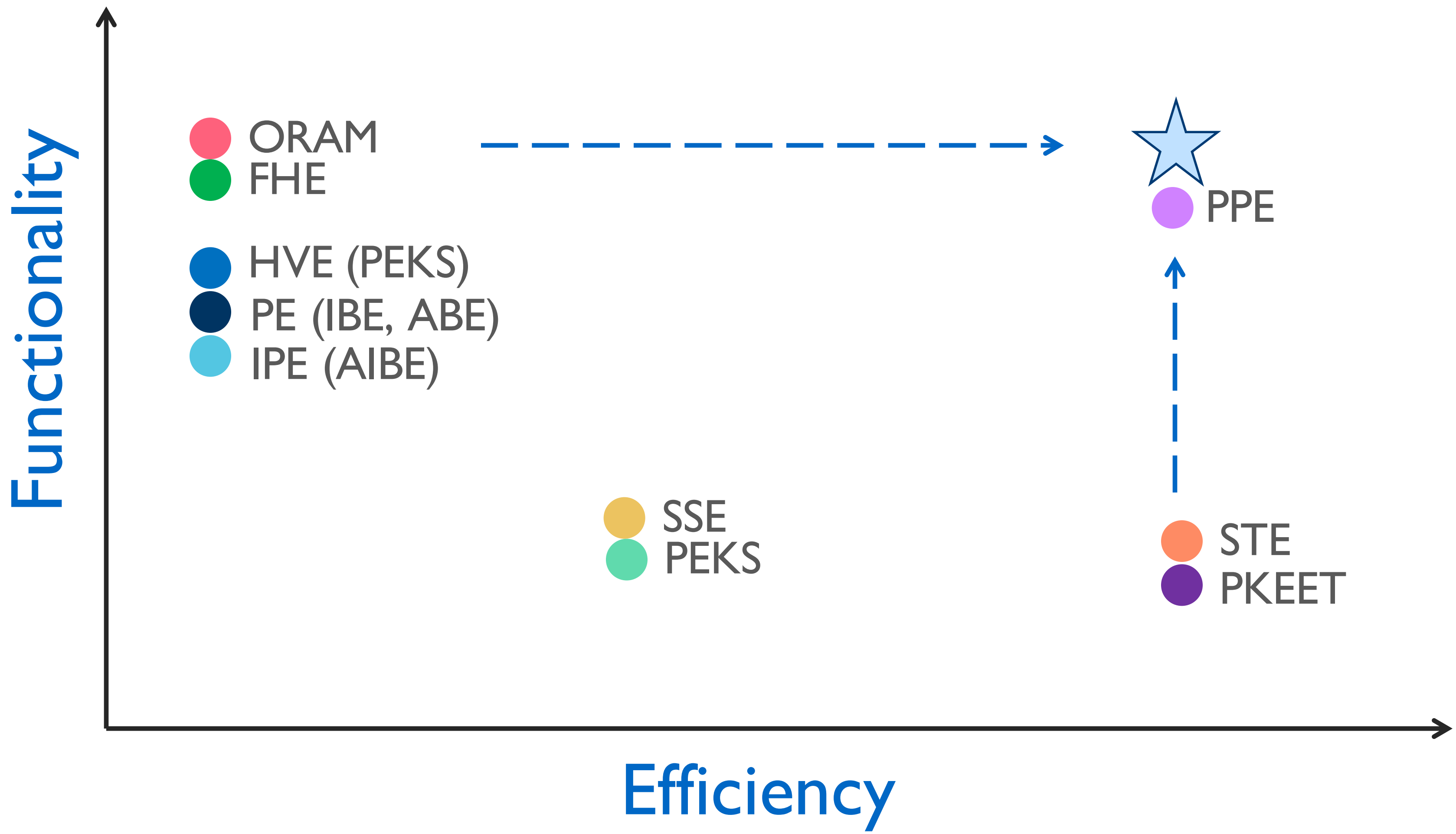
Query expressiveness
E.g. Equality, conjunctive, extended search
(subset, fuzzy, range queries, inner products)

Efficiency vs. Security



NetApp

Functionality vs. Efficiency



NetApp

Applications

- Secure search
- Secure storage
  - Outsourced, Backup
- Secure Data management
  - Deduplication, email forwarding, etc.
- Security tiers for analytics
- Private data with "enough" privacy
  - Call logs, map queries, image search, data classification

NetApp

In Practice

1) Systems
   - CryptDB, MIT CSAIL
   - Cipherbase, Microsoft
   - Google's Encrypted BigQuery Demo
   - Microsoft SQL Server 2016 Always Encrypted

2) Implementations
   - CS2, Microsoft & UCB (2012); C++; Keyword search
   - IARPA, IBM & UCI (2013); C++; Conjunctive
   - BlindSeer, Bell Labs & Columbia (2014); Boolean
   - GRECS, Microsoft, Boston & Harvard (2015); C++; Graph
   - Clusion, Brown & Colorado (2016); Java; Boolean

NetApp

Conclusion

- Tradeoffs: Security vs. Efficiency vs. Functionality
- Unclear security model
- Not-so-good asymmetric schemes
- Limited set of (academic) implementations

- But …

This could be as big a wave as public-key crypto!

Imagine a fancy animation here, in the cloud. *

Thank you.

naras@netapp.com

* You know what I mean! ☺

NetApp