Naval Information Warfare Center
PACIFIC

Can Secure Computing Solve Issues of Data Security in the Cloud?

ALLDATA 2020

Dr. Mamadou Hassimiou. Diallo, Scientist

Naval Information Warfare Center (NIWC) Pacific

San Diego, CA

U.S. Department of Defense

Naval Information Warfare Center
PACIFIC

Naval Information Warfare Center (NIWC) Pacific - Mission

Information Dominance through research, development, delivery, and support of integrated C4ISR, cyber, and space systems across all warfighting domains

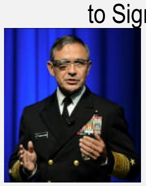
Naval Information Warfare Center PACIFIC

NIWC Focus Areas

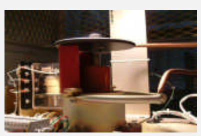
Future	Tomorrow	Today
Basic and Applied Research	Engineering, Development, Test and Evaluation	Installation and Support

Maritime Tactical C2 (MTC2) with C2 Rapid Prototyping Continuum (C2RPC) Capability

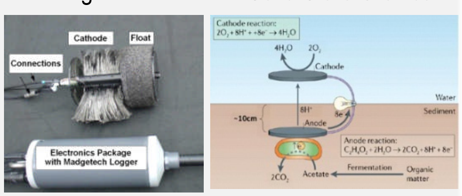
Application of Cryogenic Cooling to Signals Exploitation



Google Glass



Solid State Crowbar

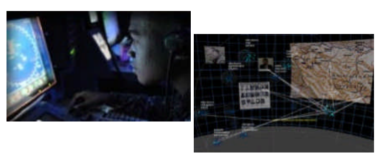


Microbial Fuel Cells: Free Power from the Mud!

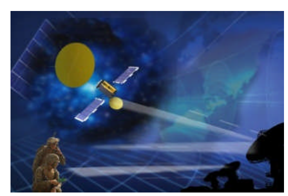
C4ISR for Unmanned Vehicles

Consolidated Afloat Networks Enterprise Services (CANES)

Ballistic Missile Defense System (BMDS)



Integrated Cyber Operations



Mobile User Objective System (MUOS)

Production, Installation and In-Service Support

Restoration and Repair

In-Service Support

Naval Information Warfare Center PACIFIC

NIWC Global Presence

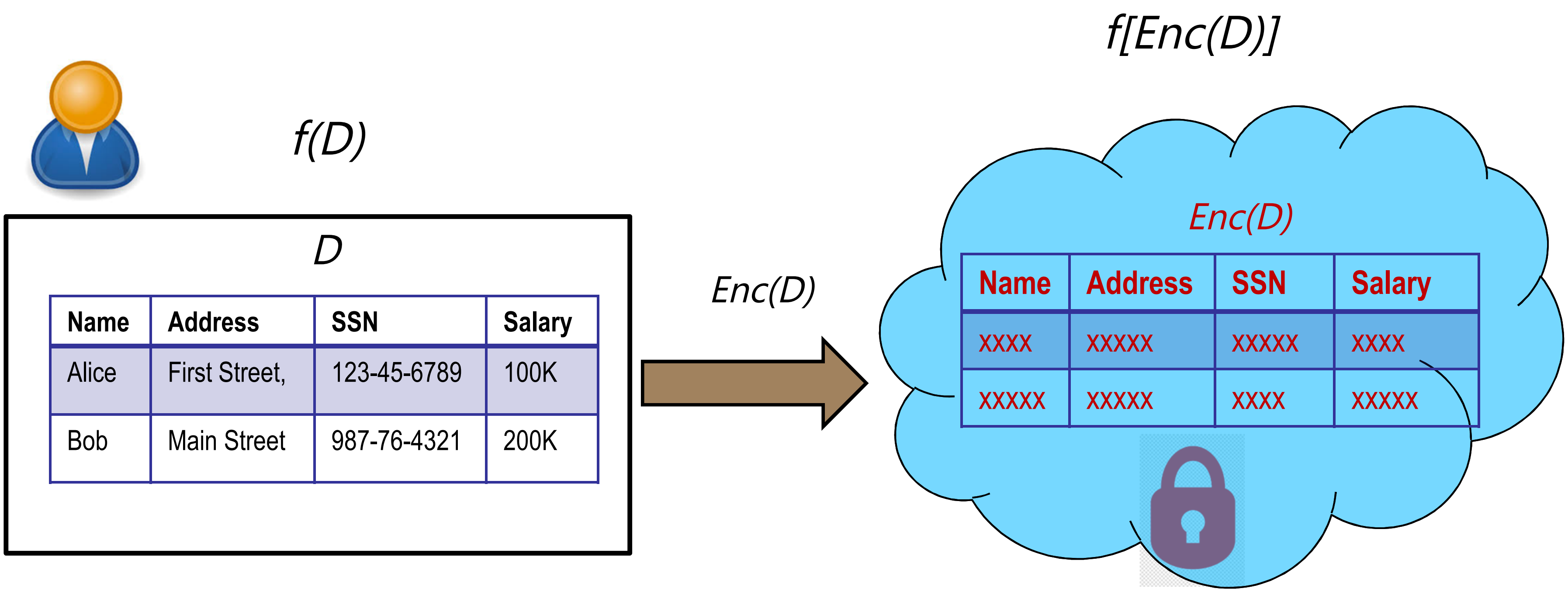
Naval Information Warfare Center PACIFIC

Can Secure Computing Solve the Issue of Data Security in the Cloud?

Secure Computing?

Homomorphic Encryption (HE) and
Multi Party Computation (MPC)

Example: Data outsourcing



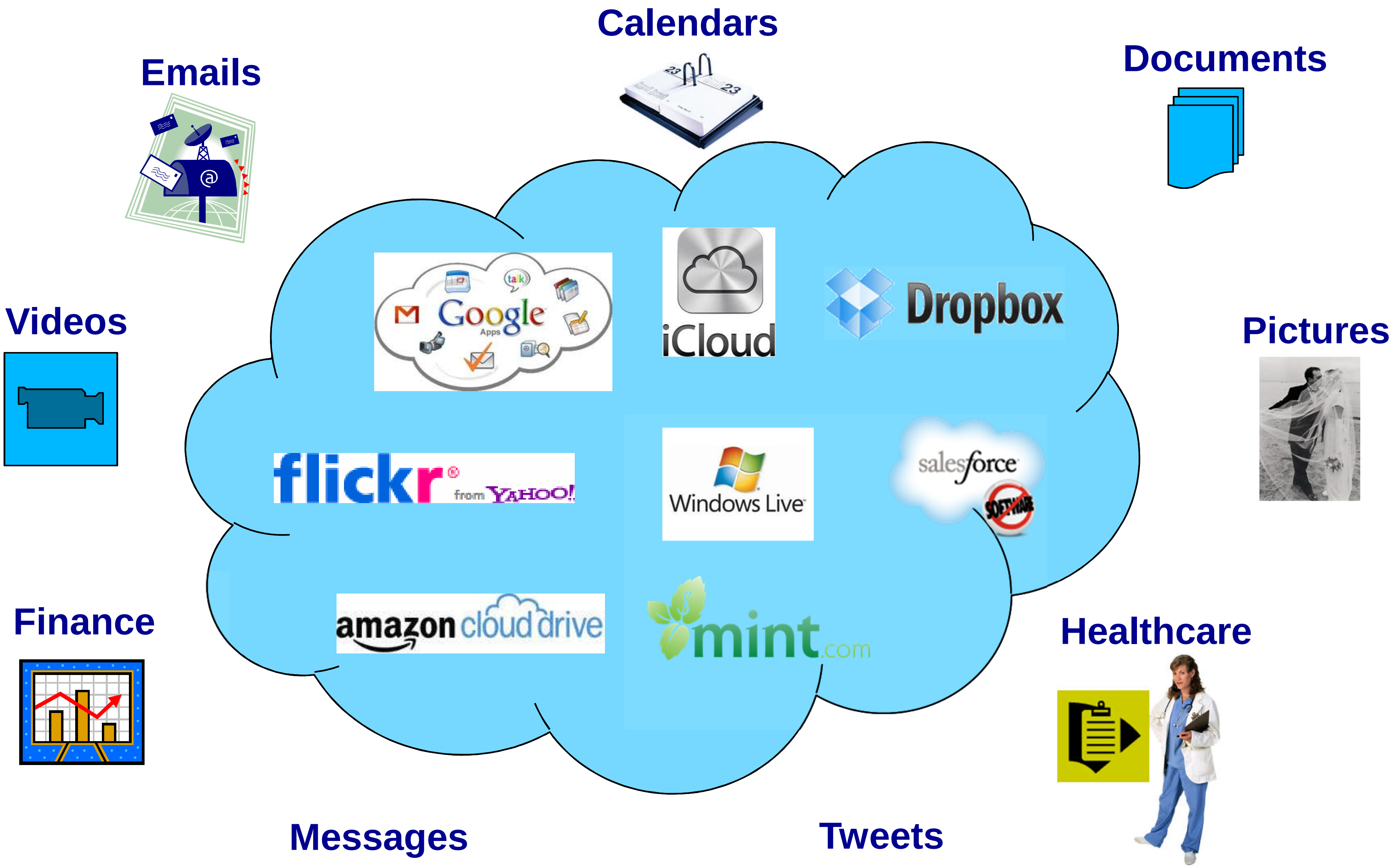
Name	Address	SSN	Salary
Alice	First Street,	123-45-6789	100K
Bob	Main Street	987-76-4321	200K

Name	Address	SSN	Salary
XXXX	XXXXX	XXXXX	XXXX
XXXXX	XXXXX	XXXX	XXXXX

Trusted Environment **Untrusted Environment**

Naval Information Warfare Center PACIFIC

Personal and Organizational Data Outsourcing



Naval Information Warfare Center PACIFIC

Data Created on the Internet Each Day

Company	Data Volume
Google	Over 300,000 Billion searches conducted worldwide daily.
Facebook	Over 4.3 Billion messages posted daily.
Twitter	Over 474,000 Tweets PER MINUTE. 682 million tweets daily.
YouTube	Over 300 hours of video are uploaded every minute. Over 4 million hours daily,
Emails	Over 293 billion emails are sent daily.
Instagram	Over 100 million photos and videos uploaded daily. Over 67,305,600 posts uploaded daily.
SMS and in-app messages	Over 100 million messages are sent every minute.

Internet Users:2014: 2.4 billion , 2016: 3.4 billion, 2017: 3.7 billion, June 2019: 4.4 billion

As the amount of data is growing so do the cyber attacks on the data.

Data Security and Privacy Issues

Personal data of 77 million users was leaked - 2011

2.9 million accounts was stolen - 2013

Data from 110 million customers was hijacked - 2013

143 million American, Canadian and British customers - 2017

data leak exposes 123 million households

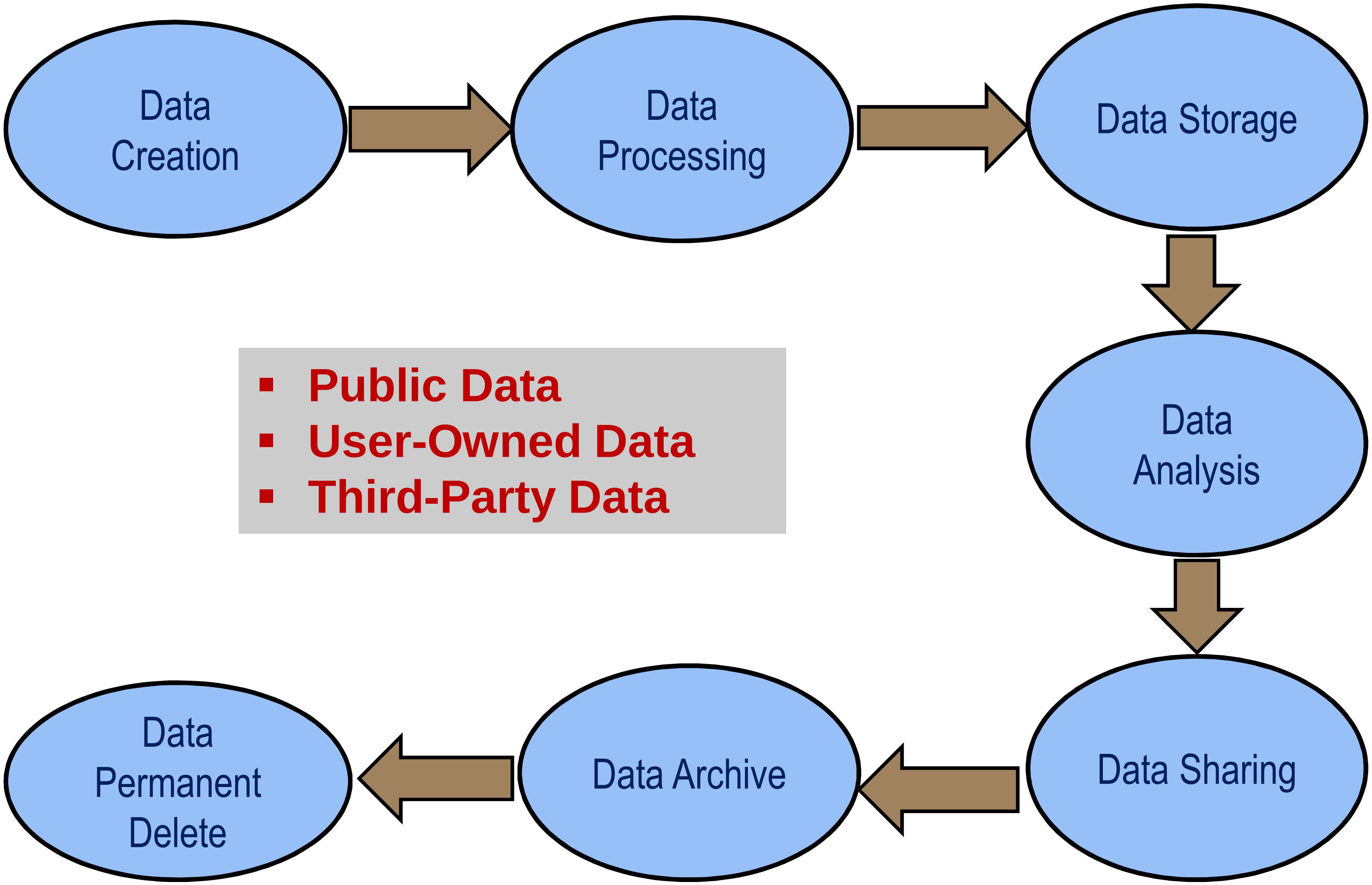
US Office of Personal Management
Records for more than 21.5 million people were stolen

WannaCry ransomware attack in May 2017
- Microsoft Windows OS
- Affected more than 200,000 computers across 150 countries

Why hackers are succeeding in stealing our data?

Naval Information Warfare Center PACIFIC

Data Lifecycle

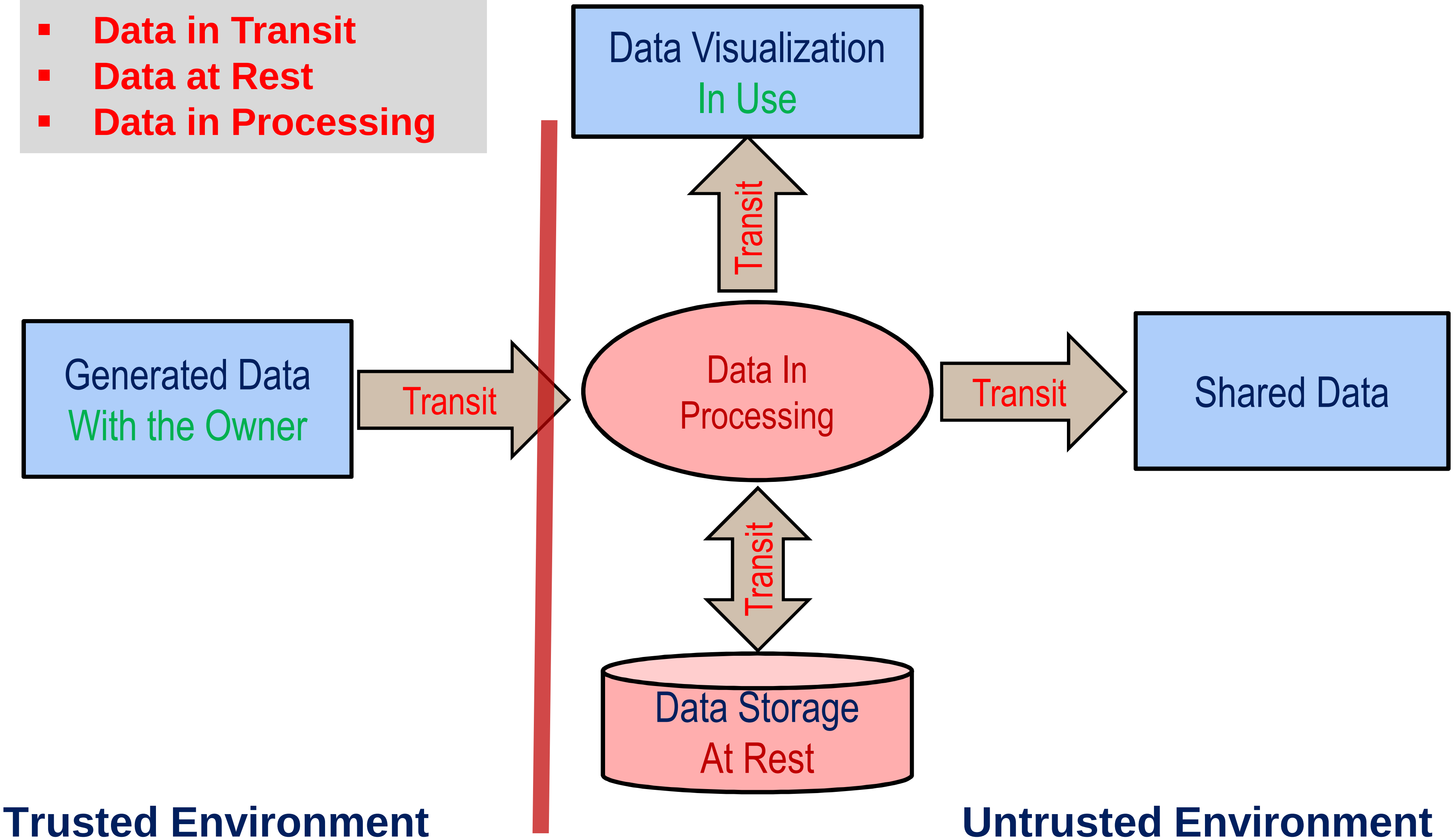


- **Public Data**
- **User-Owned Data**
- **Third-Party Data**

Naval Information Warfare Center PACIFIC

Data Lifecycle - States

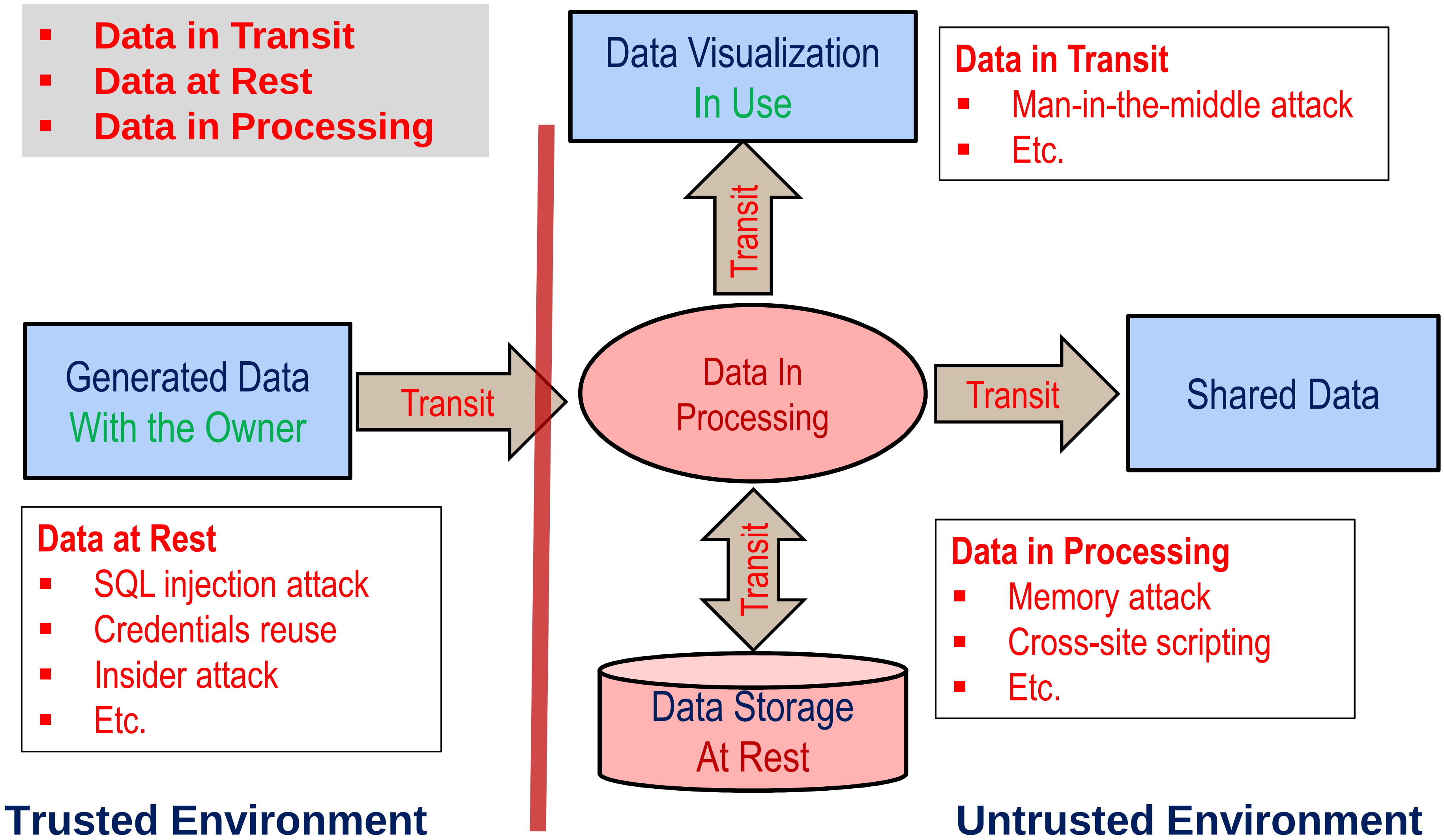
- **Data in Transit**
- **Data at Rest**
- **Data in Processing**



Trusted Environment

Untrusted Environment

Data Lifecycle – Data Security Risks



Data Security Techniques

Deterministic Encryption

Probabilistic Encryption

Data Security

Encryption 1: pSDfCAjthEsEUlPdSfn7Xw==

Encryption 2: pSDfCAjthEsEUlPdSfn7Xw==

- Not semantically secure
- Supports computation, but less secure
- RSA (No padding), Cryptographic Hash Algorithms: MD5, SHA-3, DSA

Encryption 1: pSDfCAjthEsEUlPdS5fn7Xwert

Encryption 2: eRTyuMmmZxnNn45nyhnbYuN

- Semantically secure
- Doesn't support computation, but very secure
- RSA with padding, Advanced Encryption Standard (AES), Blowfish, Serpent

Searchable Symmetric Encryption

Name	Address	SSN
Alice	First Street,	123-45-6789
Bob	Main Street	987-76-4321

Deterministic Encryption Only

Name	Address	SSN
H(Alice)	H(First) H(Street)	H(123-45-6789)
H(Bob)	H(Main Street)	H(987-76-4321)

Combination of Deterministic and Probabilistic Encryptions

Index	Name	Address	SSN
H(Alice), H(First), H(Street)	Enc(Alice)	Enc(First Street)	Enc(123-45-6789)
H(Bob), H(Main), H(Street)	Enc(Bob)	Enc(Main Street)	Enc(987-76-4321)

Leakages
- known ciphertexts – statistical analysis of messages

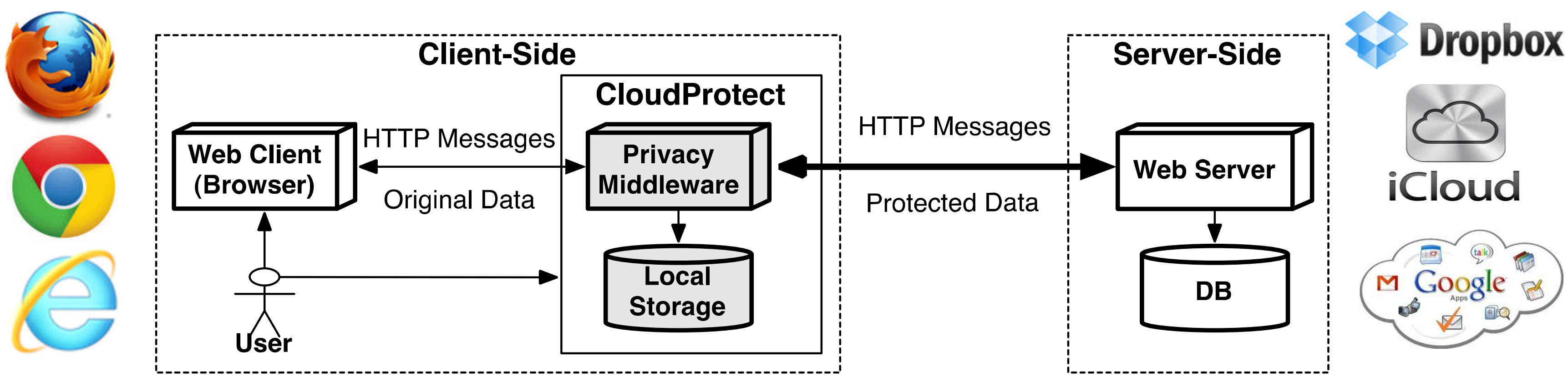
Name	Efficiency	Security
Deterministic	More	Less
Combination	Less	More

Naval Information Warfare Center PACIFIC

CloudProtect: Data Protection

Empower users to encrypt (sensitive) data within cloud applications in a transparent manner

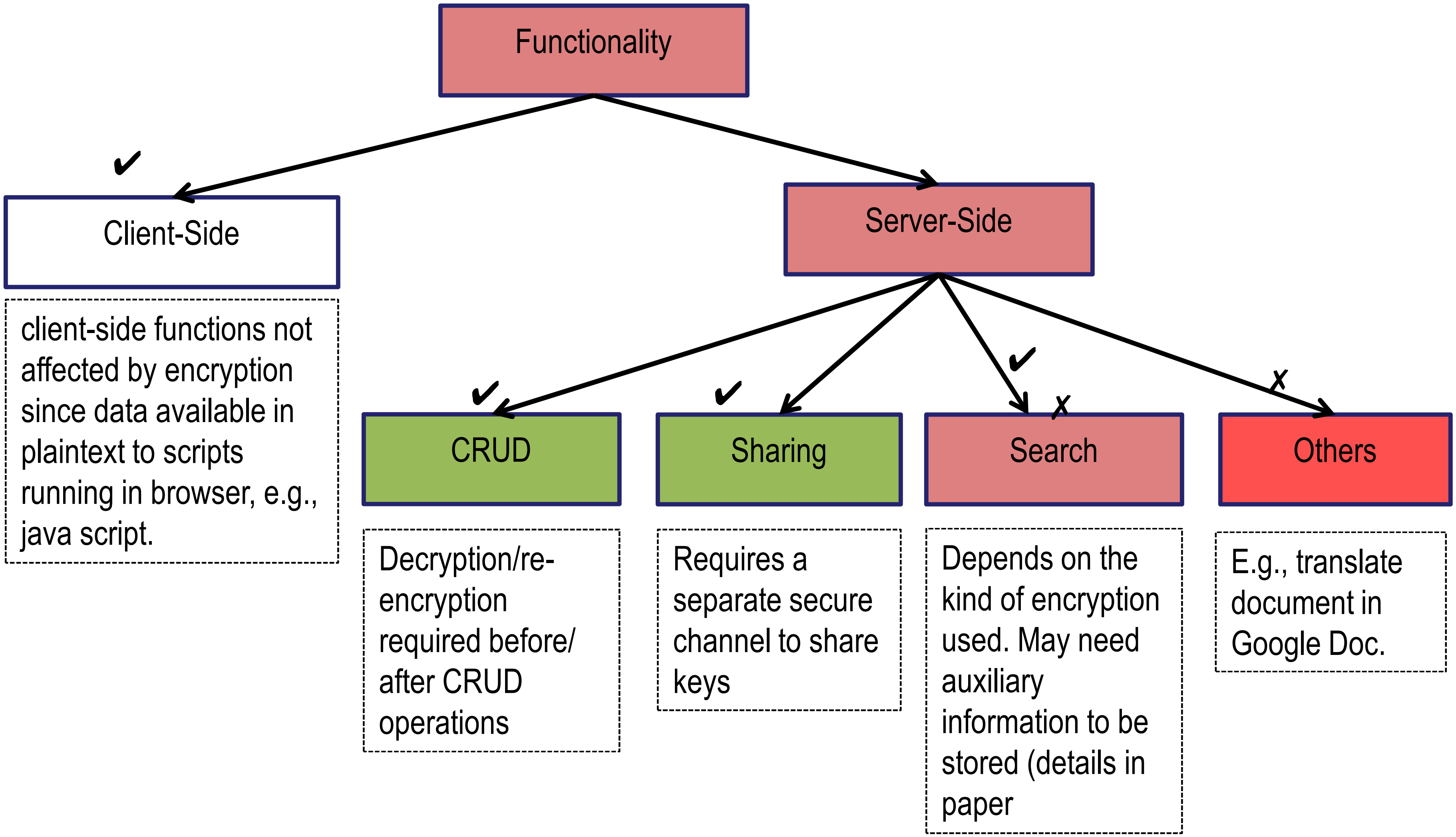
High-Level Architecture



- ***No changes to*** **the third-party server *legacy implementations***
- **Users continue to have *full access to all functionalities* offered by cloud applications.**
- ***Minimal impact on user-experience*** **by automatically adjusting what is encrypted**

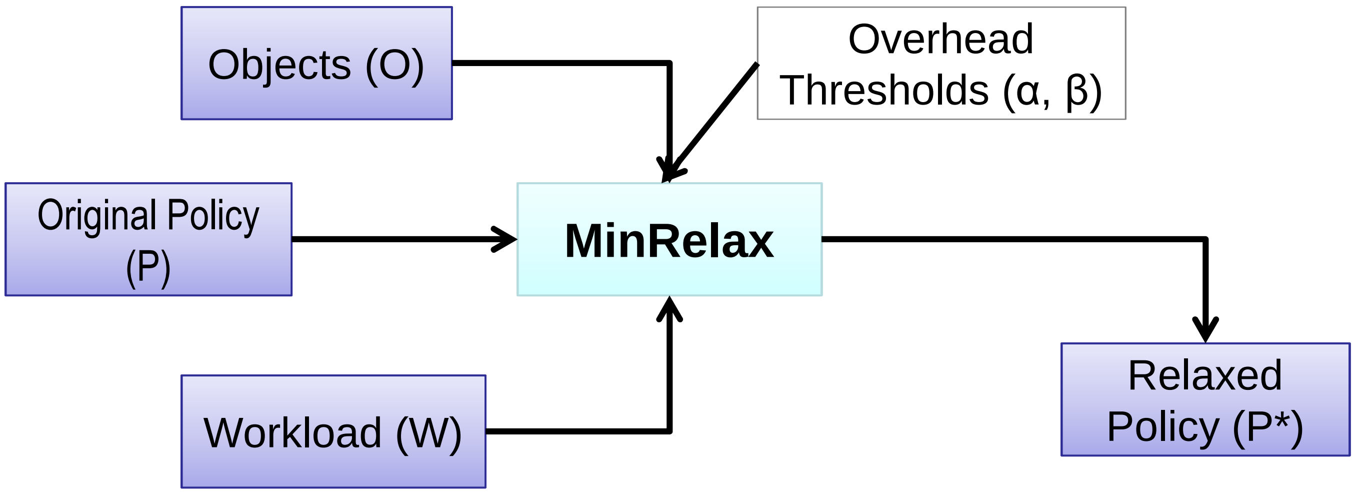
Naval Information Warfare Center PACIFIC

Computing Functions over Encrypted Representation



Optimal Tradeoff Problem

Goal: ***Reduce interruptions and/or costs while minimizing privacy reduction***



$$MinRelax(W,P,O) = ArgMin_{P^*}\left(PrivRed\left(P,P^*,O,W\right)\right)$$

Such that

$$\text{int}-oh_{avg}\left(W,P^*,O\right) \leq \alpha$$

$$\text{cost}-oh_{avg}\left(W,P^*,O\right) \leq \beta$$

NP-hard by reduction from Budgeted Max. Coverage

Naval Information Warfare Center PACIFIC

Modern Secure Computing

Secure Computing

Homomorphic Encryption

- General approach that enables computation to be performed on encrypted data
- Security model depends on cryptographic keys

- Main Issue: Computation and storage Costs

Secure Multi Party Computing

- Function over their inputs while keeping those inputs private
- Security model depends on having a network of non-colluding computers

- Main Issue: Communication cost and non-collusion requirement

Homomorphic Encryption (HE)

4	Encrypt ⟹	a87ff679a2f3e71d9181a67b7542122c
		X
5	Encrypt ⟹	e4da3b7fbbce2345d7772b0674a318d5
		=
20	⟸ Decrypt	98f13708210194c475687be6106a3b84

Types of HE Schemes

- Partially Homomorphic Encryption (PHE) – one type of gates
- Somewhat Homomorphic Encryption (SHE) - two types of gates
- Leveled Fully Homomorphic Encryption (LFHE) – arbitrary circuits of bounded depth
- Fully Homomorphic Encryption (FHE) - arbitrary circuits of unbounded depth

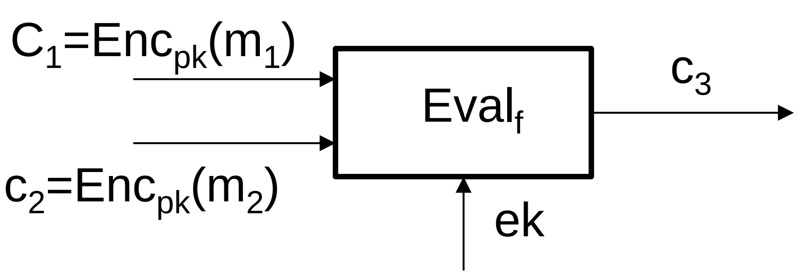
Partially Homomorphic Encryption

- **RSA**: unbounded number of modular multiplications
- **ElGamal**: unbounded number of modular multiplications
- **Goldwasser-Micali**: unbounded number of exclusive OR operations

Naval Information Warfare Center PACIFIC

Fully Homomorphic Encryption (FHE)

- Proposed by Rivest, Adleman, and Dertouzos in 1978
- First construction of FHE proposed in 2009 by Graig Gentry
  - Lattice-based cryptography
  - Computations are represented as either Boolean or arithmetic circuits with gates
  - Noisy ciphertext that grows



$Dec_{sk}(Eval_f(ek,c_1,c_2)=f(m_1,m_2)$

FHE Encryption Scheme

- “**Homomorphic**”: a (secret) mapping from plaintext space to ciphertext space that preserves arithmetic operations.
- **Mathematical Hardness**: (Ring) Learning with Errors Assumption; every image (ciphertext) of this mapping looks uniformly random in range (ciphertext space).
- “**Security Level**”: hardness of inverting this mapping without the secret key.
  - Example: 128 bits $\rightarrow 2^{128}$ operations to break
- **Plaintext**: elements and operations of polynomial ring (mod x^n+1, mod p)
  - Example: $6x^5 + 2x^4 + 3x^3 + \ldots$
- **Ciphertext**: elements and operations of polynomial ring (mod x^n+1, mod q)
  - Example: $7862x^5 + 5652x^4 + \ldots$

BGV Homomorphic Encryption Scheme

- ## The Ring Learning with Errors (LWE) Problem

For security parameter λ, let $f(x) = x^d + 1$ where $d = d(\lambda)$ is a power of 2.

Let $q = q(\lambda) \geq 2$ be an integer.

Let $R = Z[x]/(f(x))$ and let $R_q = R/qR$.

Let $\chi = \chi(\lambda)$ be a distribution over R.

The $RLWE_{d,q,\chi}$ problem is to distinguish the following two distributions:

In the first distribution, one samples (a_i, b_i) uniformly from $R^2{}_q$.

In the second distribution, one first draws $s \leftarrow R_q$ uniformly and then samples (a_i, b_i) in $R^2{}_q$ by sampling $a_i \leftarrow R_q$ uniformly, $e_i \leftarrow \chi$, and setting

$bi = a_i \cdot s + e_i$

The $RLWE_{d,q,\chi}$ assumption is that the $RLWE_{d,q,\chi}$ problem is infeasible.

BGV Homomorphic Encryption Scheme

- **Preliminaries**
  - Polynomial rings: $A = Z[X]/\Phi_m(X)$,
    where m is parameter and $\Phi_m(X)$ m'th cyclotomic polynomial
  - Native plaintext space: Ring: $A_2 = A/2A$, binary polynomials modulo $\Phi_m(X)$
  - Ciphertext space: vectors over $A_q = A/qA$, q an odd modulo
  - Chain of modulus: $q_0 < q_1 < \text{---} < q_L$
  - Level-i ciphertexts: $c = (c_0, c_1) \in (A_{qi})^2$, 2-element vectors over R_{qi}
- **Definition**
  - Secret keys: $s \in A$, and $s = (1, s)$, with small coefficients
  - Level-i cipher $c = (c_0, c_1)$ encrypts a plaintext polynomial $m \in A_2$,
    with respect to $s = (1, s)$ if we have the equality over A,
    $[\langle c, s\rangle]_{qi} = [c_0 + s.c_1]_{qi} \cong m \pmod 2$
  - Noise: $[c_0+s.c_1]_{qi}$, is small

FHE Schemes and Libraries

Scheme	Year	Security
Brakerski-Gentry-Vaikuntanathan (BGV)	2011	Ring Learning With Errors (RLWE) problem
NTRU-based scheme by Lopez-Alt, Tromer, and Vaikuntanathan (LTV)	2012	Variant of the NTRU computational problem
Brakerski/Fan-Vercauteren (BFV)	2012	Ring Learning With Errors (RLWE) problem
NTRU-based scheme by Bos, Lauter, Loftus, and Naehrig (BLLN)	2013	Variant of the NTRU computational problem
Craig Gentry, Amit Sahai, and Brent Waters (GSW)	2013	Ring Learning With Errors (RLWE) problem
FHEW	2014	Ring Learning With Errors (RLWE) problem
TFHE	2016	Ring Learning With Errors (RLWE) problem
The Cheon-Kim-Kim-Song (CKKS)	2016	Ring Learning With Errors (RLWE) problem

FHE Schemes and Libraries

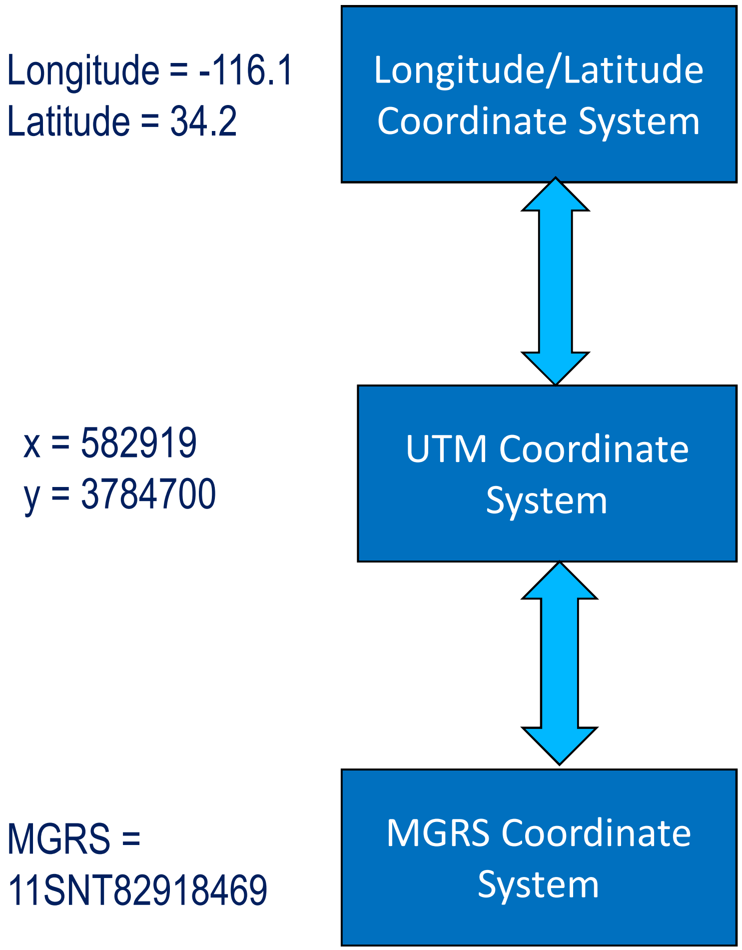
Library/Scheme	FHEW	TFHE	BGV	BFV	CKKS
cuFHE		✔			
FHEW	✔				
FV-NFLlib				✔	
HEAAN					✔
HElib			✔		(✔)
PALISADE			✔	✔	(✔)
SEAL				✔	✔
TFHE(-Chimera)	✔	✔		(✔)	(✔)

CallForFire

- **A Call For Indirect Fire Support is a protocol for destroying enemy targets**
- **Players**
  - Forward Observer (FO): locates High Value Target
  - Fire Direction Center (FDC): command & control
  - Firing Unit (FU): shoots enemy target
  - High Value Target (HVT): enemy target
  - Observer-Target (OT): distance between the observer and the target
- **Mission-critical system which requires high security & privacy**
  - Forward Observer is at risk
  - Insider threat

Naval Information Warfare Center PACIFIC

CallForFire



- **Storing Location**
  - Location attributes stored as key/value pairs $<hash(id_{location}+label), Enc(Value)>$
  - Sine and Cosine shifted to get integers
  - Easting and Northing also shifted

Hash(id+label)	Enc(Value)
Hash(101+id)	Enc(101)
Hash(101+Easting)	Enc(829100)
Hash(101+Northing)	Enc(846900)
Hash(101+distance)	Enc(2000)
Hash(101+Sine)	Enc(87)
Hash(101+Cosine)	Enc(50)
…	…

CallForFire

- **Computing HVT Location**
  - Inputs: FO easting and FO northing, OT (distance between FO and HVT) Θ (angle of HVT relative to due North)
  - Outputs: HVT easting, HVT northing

$$\mathrm{HVT}_{\mathrm{easting}} = \mathrm{FO}_{\mathrm{easting}} + \mathrm{OT}_{\mathrm{distance}} \times \sin(\theta)$$

$$\mathrm{HVT}_{\mathrm{northing}} = \mathrm{FO}_{\mathrm{northing}} + \mathrm{OT}_{\mathrm{distance}} \times \cos(\theta)$$

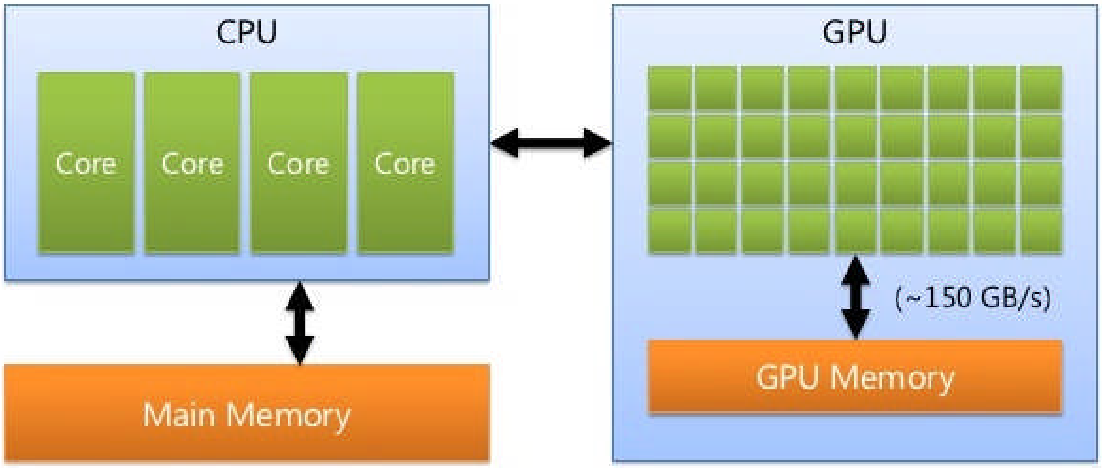
- **Computing Distance FU to HVT**
  - Inputs: FU and HVT eastings and northings
  - Inputs: Dist^2 (distance between FU and HVT)

$$\mathrm{Dist}^2_{\mathrm{FU\text{-}HVT}} = (\mathrm{FU}_{\mathrm{easting}} - HVT_{\mathrm{easting}})^2 + (\mathrm{FU}_{\mathrm{northing}} - HVT_{\mathrm{northing}})^2$$

Naval Information Warfare Center PACIFIC

CallForFire - GPGPU

- GPGPU board has 1,000's of cores
- Massive parallelism can be exploited
- GPU used for parallelizing HElib
- Profiling decision based on the following formula:
- NVIDIA CUDA library used for implementation



$$T_{CPU} > T_{GPU} =$$
$$T_{GPU\_init} + T_{GPU\_runtime} + T_{CPU \rightarrow GPU}$$
$$+ T_{CPU \leftarrow GPU} + T_{mem\_alloc} + T_{mem\_dealloc}$$

Naval Information Warfare Center PACIFIC

CallForFire - GPGPU

- TAU Parallel Performance System used to gather timing information
- FHE Setup is costly
- GPU implementation is faster
- Parallelized Algorithm: Bluestein's Fast Fourier Transform (FFT)
- FFT computes Discrete FT

Table VI: Comparison of CPU and GPU BluesteinInit/FFT Implementation Combinations (256 Threads Per Block)

	CUDA Overhead (ms)	Workload Exec. (ms)	Total Exec.
CPU BluesteinInit(), CPU BluesteinFFT()	0	3,970	3,970
GPU BluesteinInit(), CPU BluesteinFFT()	56	3,836	3,892
CPU BluesteinInit(), GPU BluesteinFFT()	46	2,119	2,166
GPU BluesteinInit(), GPU BluesteinFFT()	43	2,033	2,077

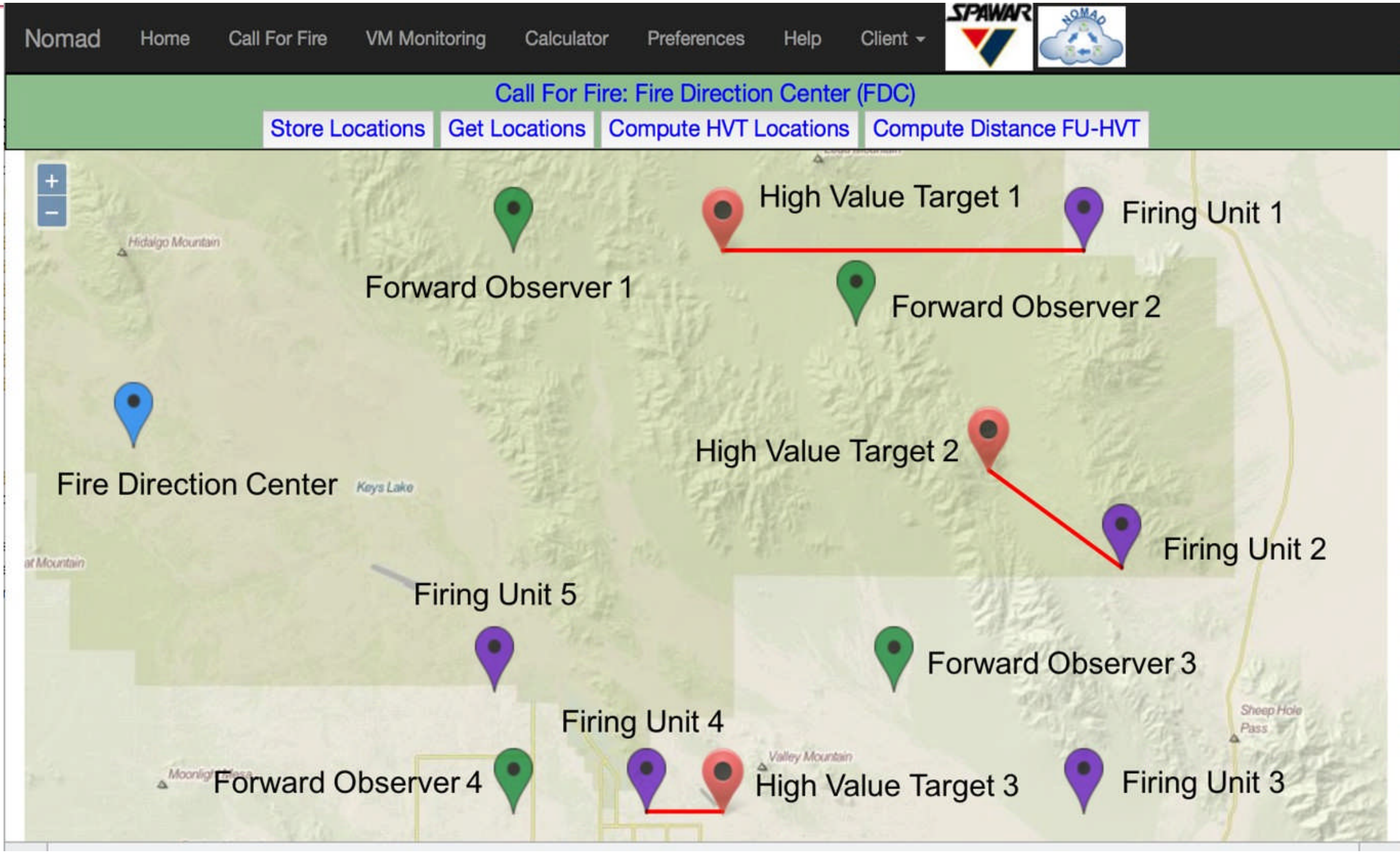
Table II: HElib Profiling Results

	CPU HELib Exec. Time (ms)	GPU HELib Exec. Time (ms)
FHE Context	1,278	1,136
Mod Chain	412	298
Secret Key	245	36
Key Switching Matrices	1,826	509
Encoding Single Value	0.108	0.114
Encrypting Single Value	65	65
Adding Two Values	1	0.465
Decrypting Result	28	8
Decode Result	0.116	0.0358

Table V: GPU Overhead

	BluesteinInit()	BluesteinFFT()	Average
GPU Init	38.561 ms	92.713 ms	80.226 ms
cudaMalloc	151.01 us	32.889 us	32.855 us
cudaMemcpy, Host to Device	NA	4.8030 us	4.8040 us
cudaMemcpy, Device to Host	6.3040 us	6.2040 us	6.2180 us
cudaFree	72.231 us	80.318 us	72.381 us
OverHead Kernel Exec.	230us/14KB	124us/14KB	116us/14KB

CallForFire Implementation



CallForFire Experiments

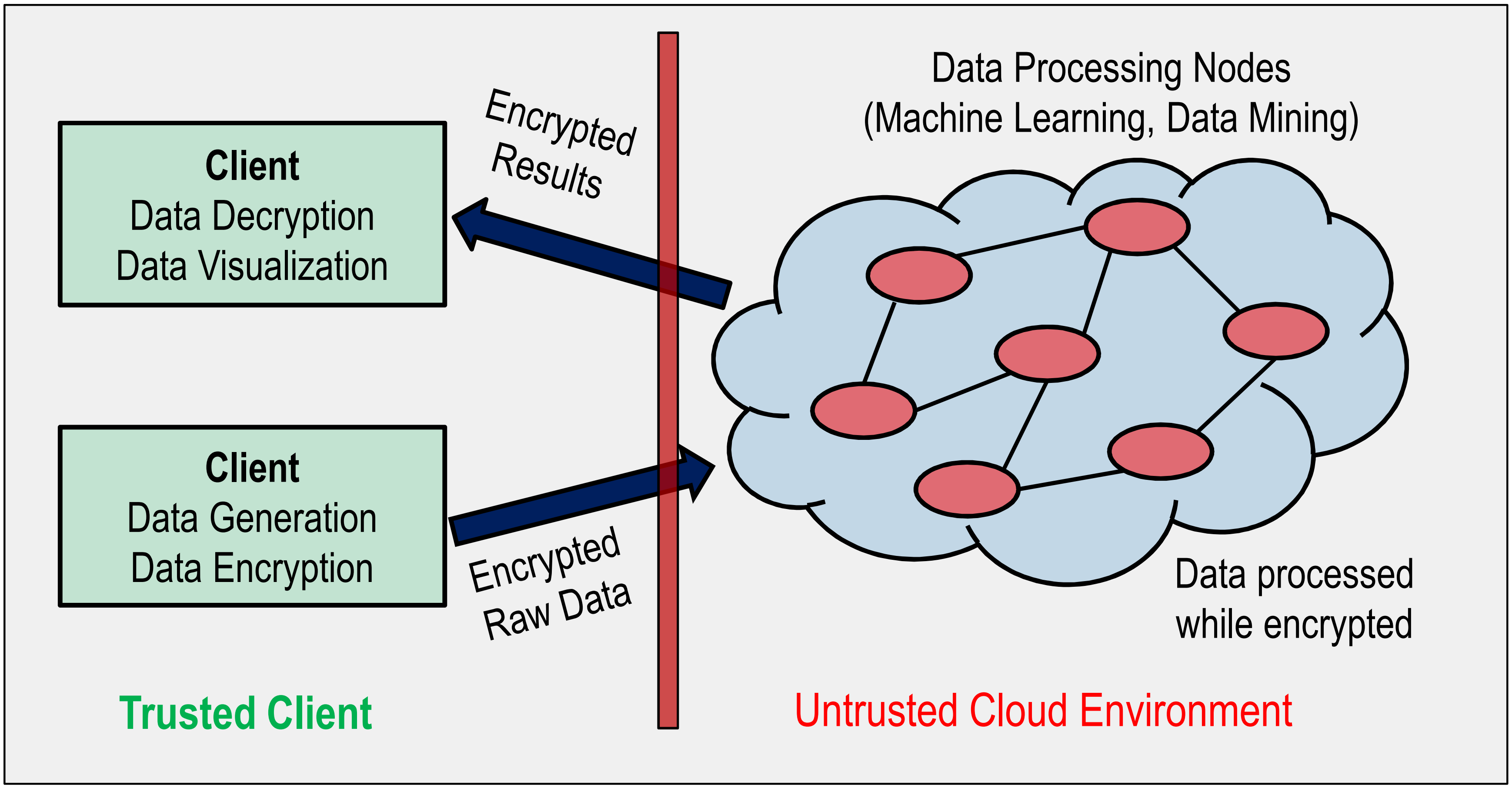
Number of FOs: 10
Number of FUs: 10
Location Computation: 1 HVT for each FO
Distance Computation: pairwise between FU and HVT

Table 1. Average Computation Overhead in Sec. with Fixed p=9576890767 (10 digits)

k: Security parameter	80 (L=11, m=11021)		100 (L=11, m=12403)		120 (L=11, m=13019)	
Type	Individual	Batched	Individual	Batched	Individual	Batched
Location Encryption	702.3990	63.0778	782.6890	71.8735	831.9190	77.1963
Location Decryption	600.7040	165.2790	692.3490	217.0520	760.9390	217.0620
Location Computation	212.1974	21.3238	221.7478	27.3559	237.1199	23.2551
Distance Computation	271.2946	26.3864	283.7557	28.7946	331.0418	33.2885
Storing Location	2.4743	0.2498	2.7999	0.2847	2.8119	0.2824
Retrieving Location	16.3833	1.5589	18.0937	1.8003	21.8311	1.9645

Distributed Framework for Secure Computation in the Cloud (DFSCC)

***DFSCC**: leverages HE to enable analytic tools to process encrypted data in a large-scale distributed system in the cloud.*



Distributed Framework for Secure Computation in the Cloud (DFSCC)

Cloud Infrastructure as a Service: Deployed using the open source Xen hypervisor

Distributed System: Apache SPARK

Homomorphic Encryption Library: PALISADE

Machine Learning Algorithm: Support Vector Machine (SVM) for classifying data

Running SVM locally under different matrix sizes

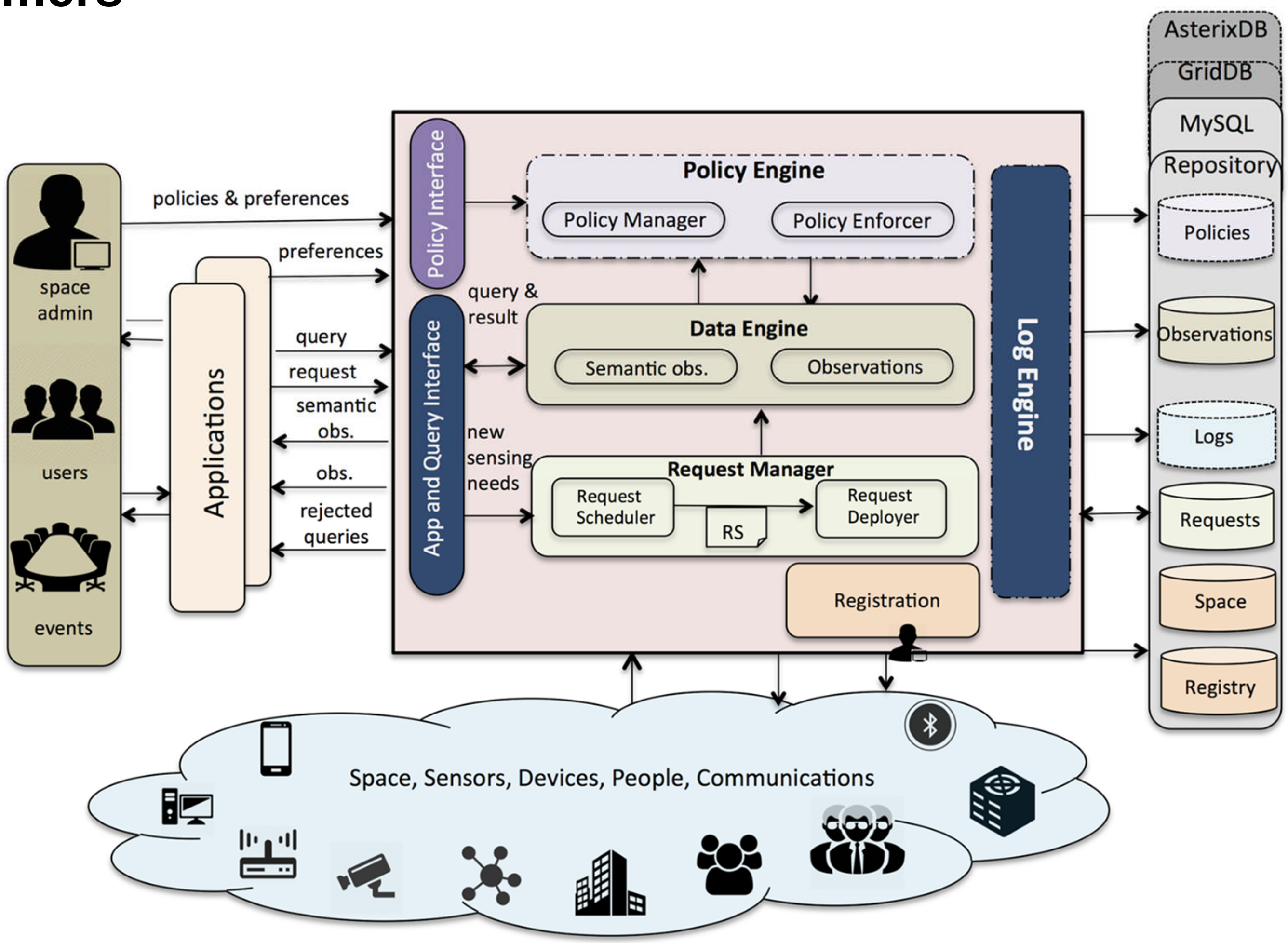
Inputs		Time (sec.)				Time (sec.)
Matrix	Vector	Encrypt	Compute (cipher)	Decrypt		Compute (plaintext)
8 x 2048	8	0.4938	0.1050	0.0497		0.00027
16 x 2048	16	0.9001	0.1282	0.0503		0.00030
32 x 2048	32	1.8193	0.1981	0.0503		0.00037
64 x 2048	64	3.6415	0.3233	0.0500		0.00065

TIPPERS (Testing of Testbed for IOT-Based Privacy Preserving Pervasive Spaces)

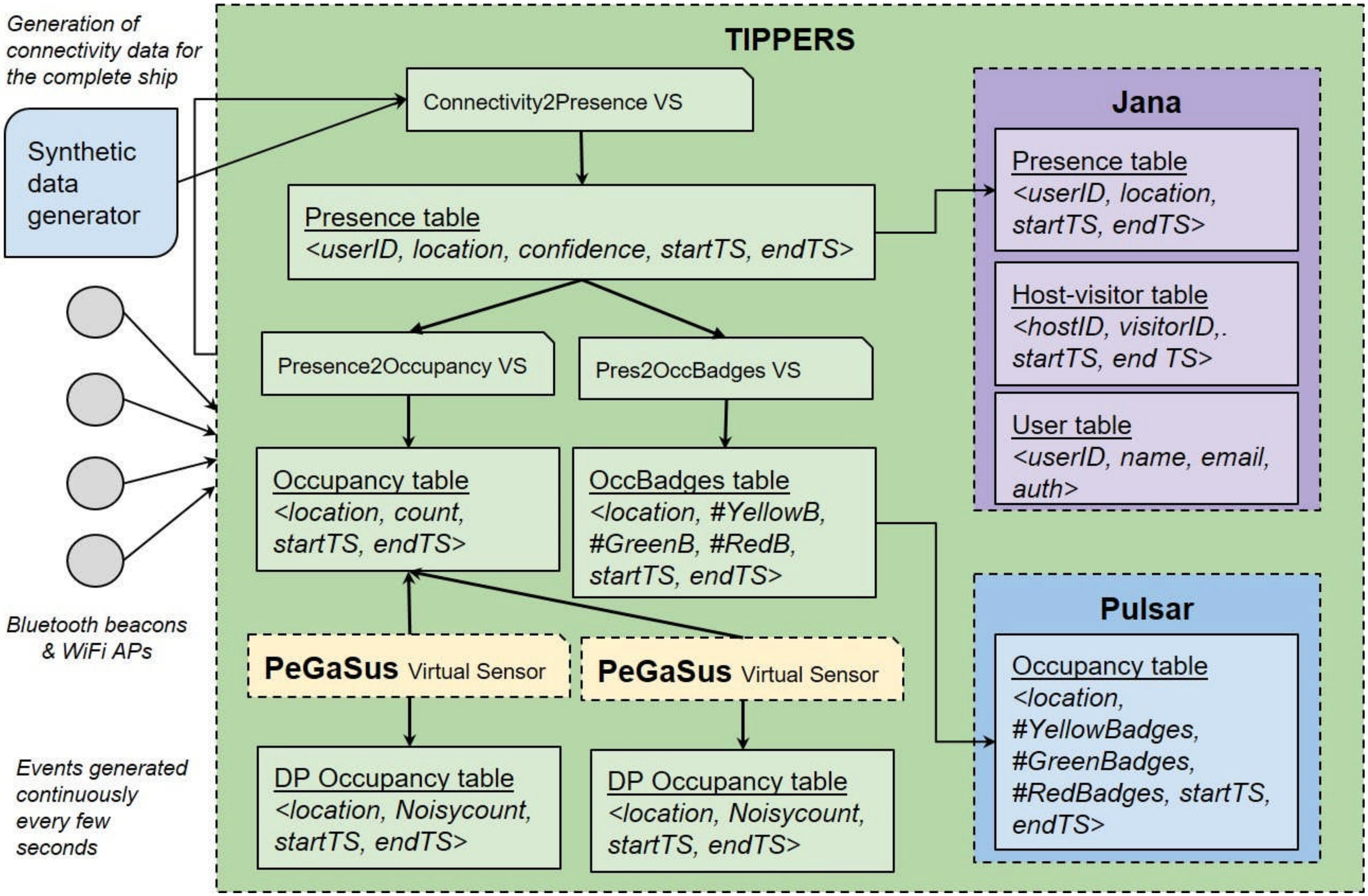
Naval Information Warfare Center PACIFIC

Funded by the DARPA Brandeis Program

9 Performers



TIPPERS – Trident Warrior 2019



TIPPERS RIMPAC/Trident Warrior 2020

Homomorphic Encryption Standardization

Goal

HomomorphicEncryption.org is an open consortium of industry, government and academia to standardize homomorphic encryption.

Current Participants

- **Industry:** Microsoft, Samsung SDS, Intel, Duality Technologies, IBM, Inpher, Google SAP, Intuit, General Dynamics, Mastercard, CryptoExperts, Algorand, Foundation, Mercedes Benz, Alibaba Group, LinkedIn, IXUP, Intertrust

- **Government:** NIH, NIST, NSF, NIWC, SLAC National Accelerator Lab, United Nations / ITU

- **Academia:** Boston University, Brown, CEA LIST, Columbia, EPFL, MIT, NJIT, NYU, Royal Holloway University, RIT, UCSD, Univ. of Cincinnati, Univ. of Hannover, Univ. of Michigan, Univ. Texas Austin, Univ. of Toronto, UC Irvine, Univ. of Waterloo, Sabanci University, Seoul National University, WPI

Summary

Can Secure Computing Solve Issues of Data Security in the Cloud?
YES

However

- It requires more patriations from academia, industry, government
- More investment is needed
- May require combination of hardware and software approaches

U.S. Department of Defense programs focusing on secure computing

- DARPA PROCEED
- DARPA Brandeis
- DARPA DRIVE
- IARPA HECTOR

Questions?

Dr. Mamadou H. Diallo, Scientist
Naval Information Warfare Center (NIWC) Pacific
U.S. Department of Defense
mamadou.h.diallo@navy.mil