Cryptographic Memory Tagging:

Towards Stateless Integrity

HASP ’24 | November 2, 2024

Bharath Namboothiry (University of Pennsylvania, Intel Labs)

David Durham, Christoph Dobraunig, Michael LeMay (Intel Labs)

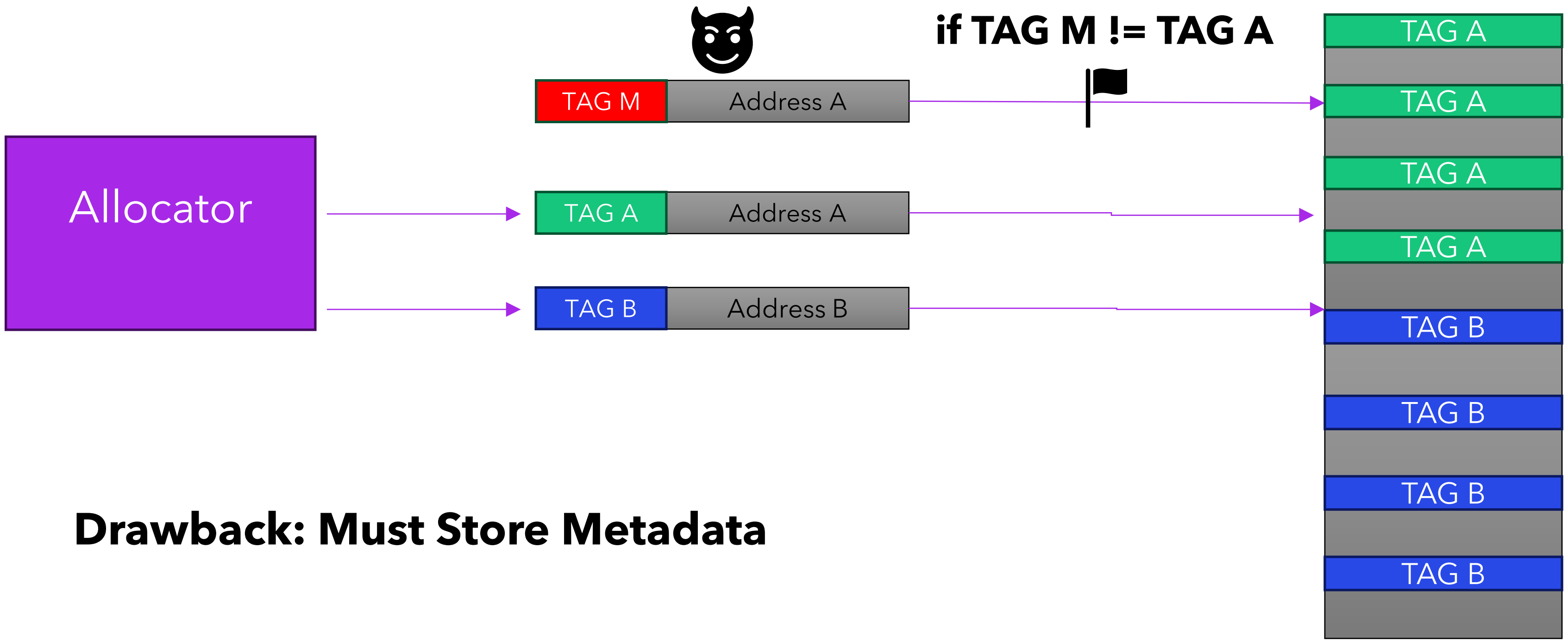
This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Project Overview

- **Memory safety violations** (Use After Free, Buffer Overflow, etc) persist, causing unauthorized access, data corruption, and system crashes
- Traditional tagging solutions **require metadata storage**, introducing significant overhead
- **Our approach** implements tagging via a cryptographic pointer framework, allowing memory access control **with near-zero metadata storage**
- Simulated results show good integrity coverage with negligible memory overhead

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

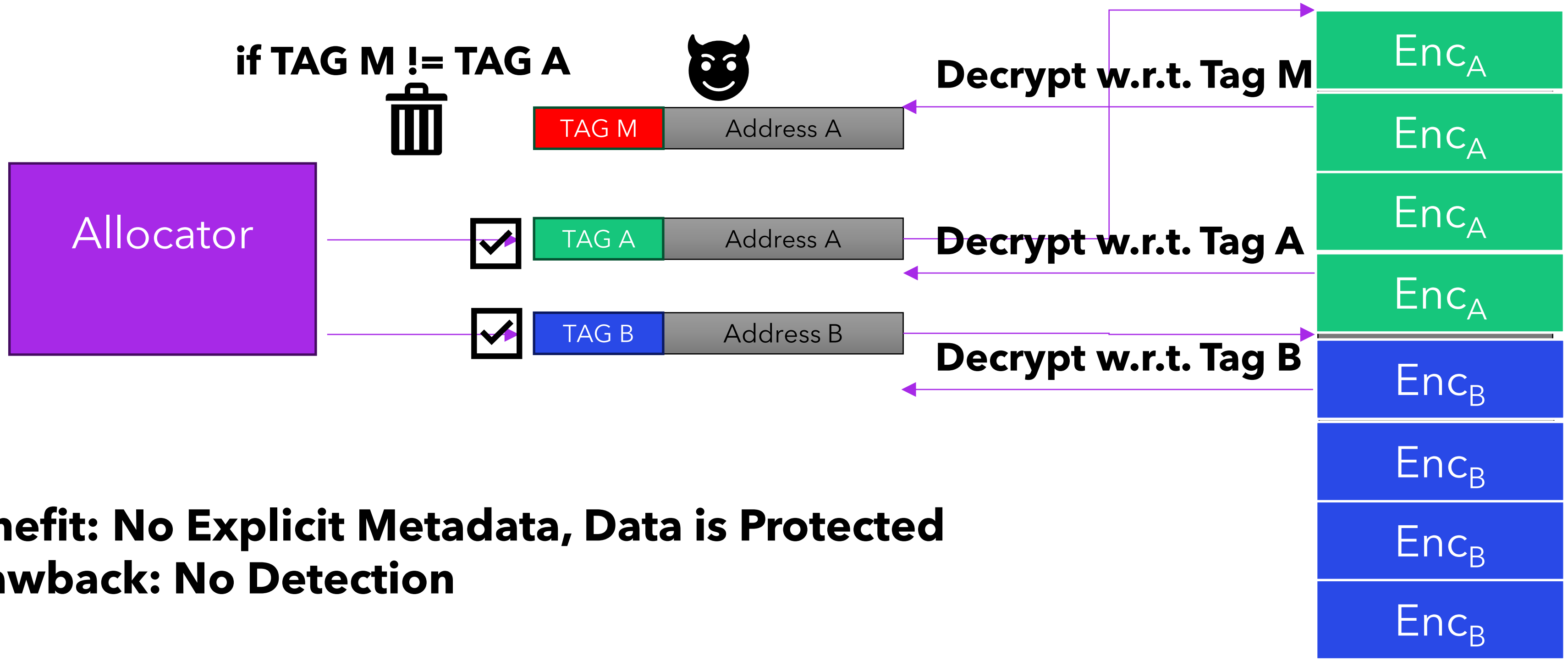
Traditional Memory Tagging



Drawback: Must Store Metadata

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Cryptographic Capability Computing (C3) Memory Safety



Benefit: No Explicit Metadata, Data is Protected
Drawback: No Detection

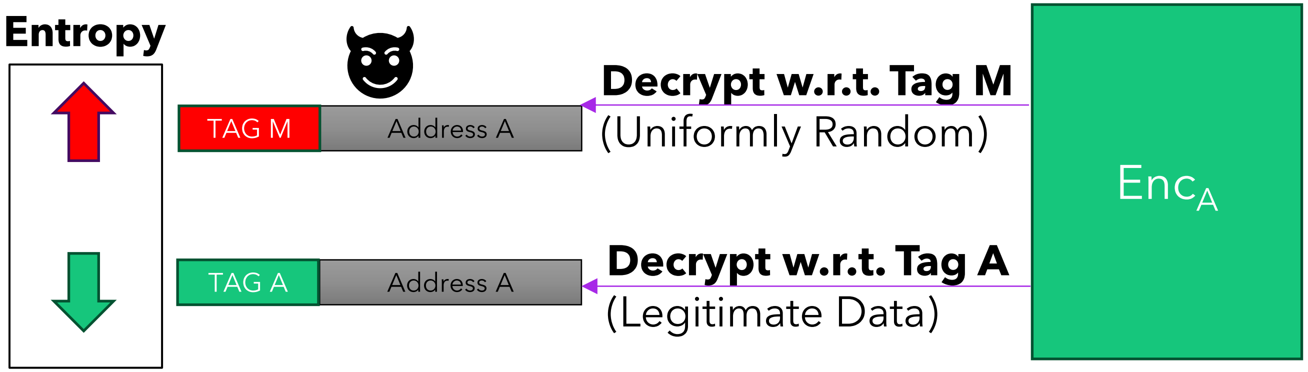
This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Cryptographic Memory Tagging

- **Goal:** Use the C3 encryption framework to perform memory tagging without state
- **Strategy:** Infer integrity from decryption entropy

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Key Insight: Decryption Entropy



This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Binary Entropy Testing

Input: Data Granule

Output: Decision Boolean (high/low Entropy)

A <u>useful</u> binary entropy test:

Uniformly Random Granule : ↑ With provably high probability

Legitimate Data Granule : ↓ For *most** workload data

Let's assume for now we have something like this on hand (see appendix slides)

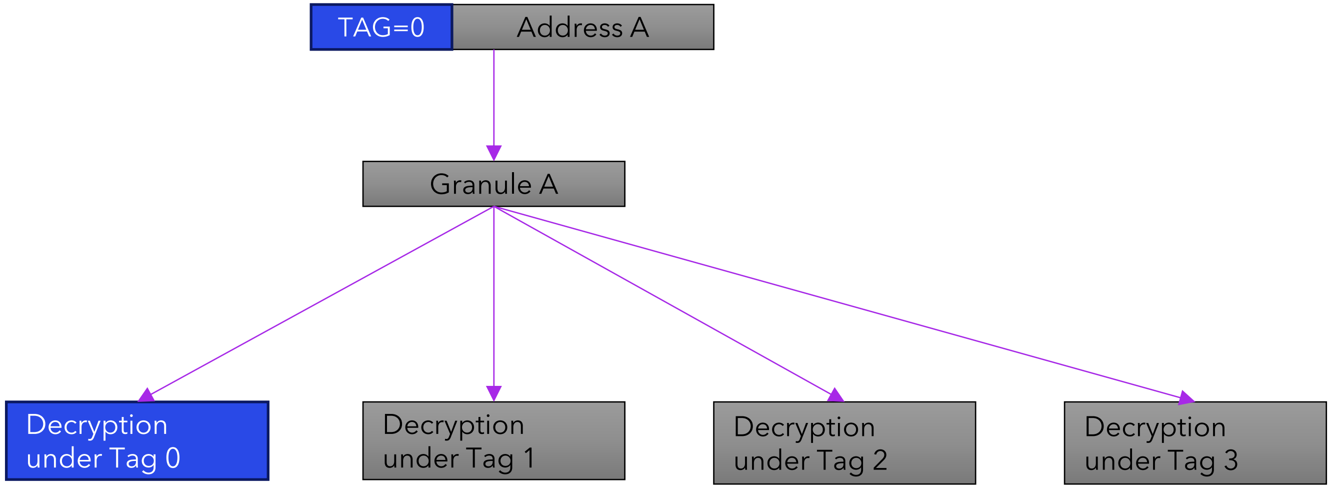
This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Inferring Integrity

Uniformly Random Granule : With provably high probability

Legitimate Data Granule : For *most** workload data

(Assume 2-bit tags)



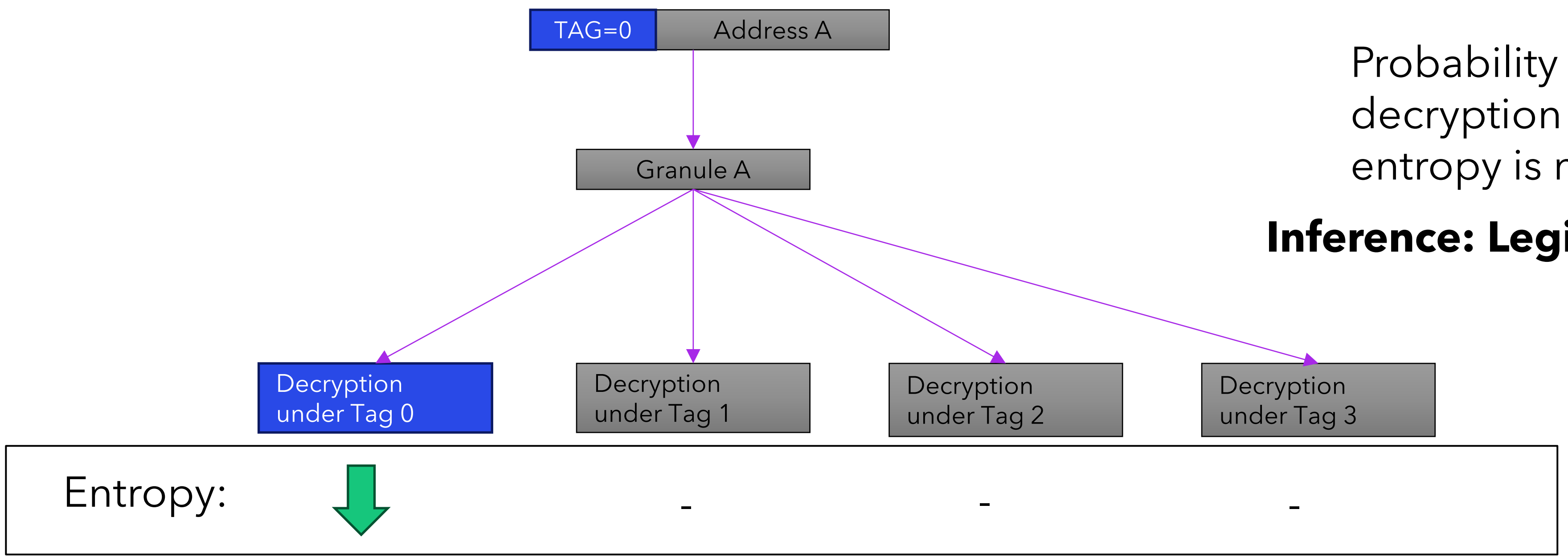
This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Inferring Integrity

Uniformly Random Granule : With provably high probability

Legitimate Data Granule : For *most** workload data

(Assume 2-bit tags)



Probability of incorrect decryption with low entropy is near-zero

Inference: Legitimate Access

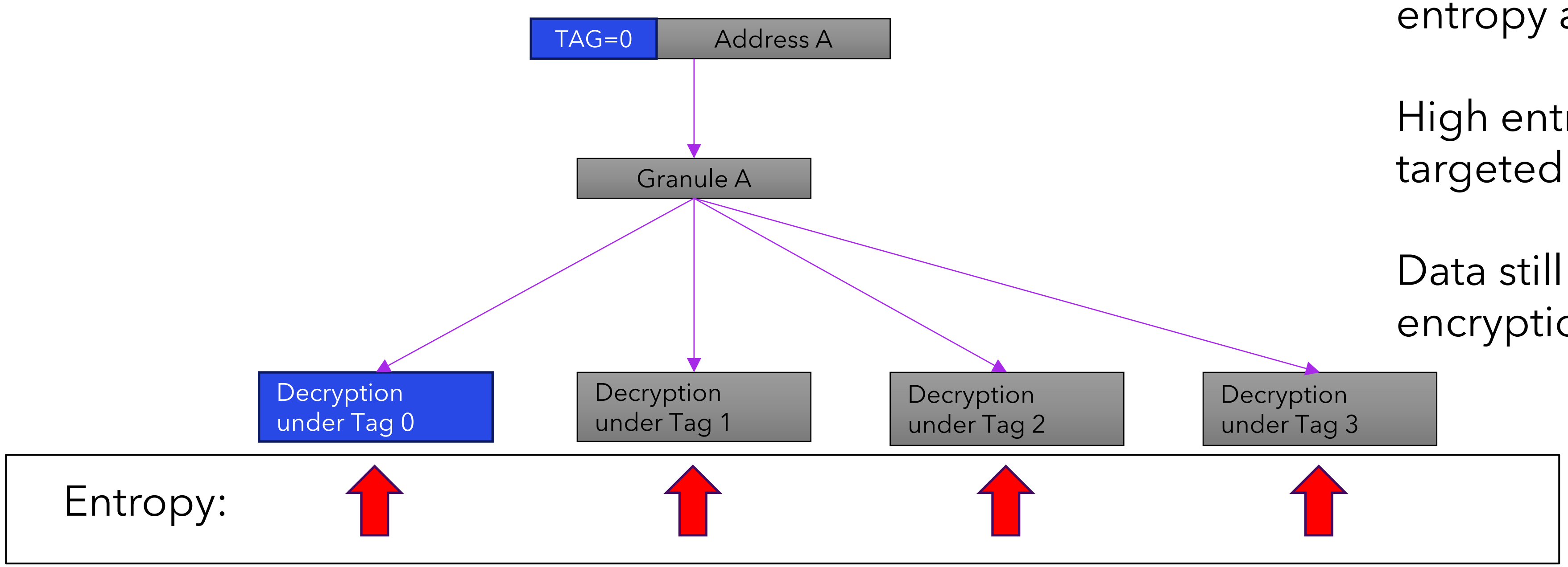
This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Inferring Integrity

Uniformly Random Granule : With provably high probability

Legitimate Data Granule : For *most** workload data

(Assume 2-bit tags)



The true decryption is high entropy and indistinguishable

High entropy data typically isn't targeted

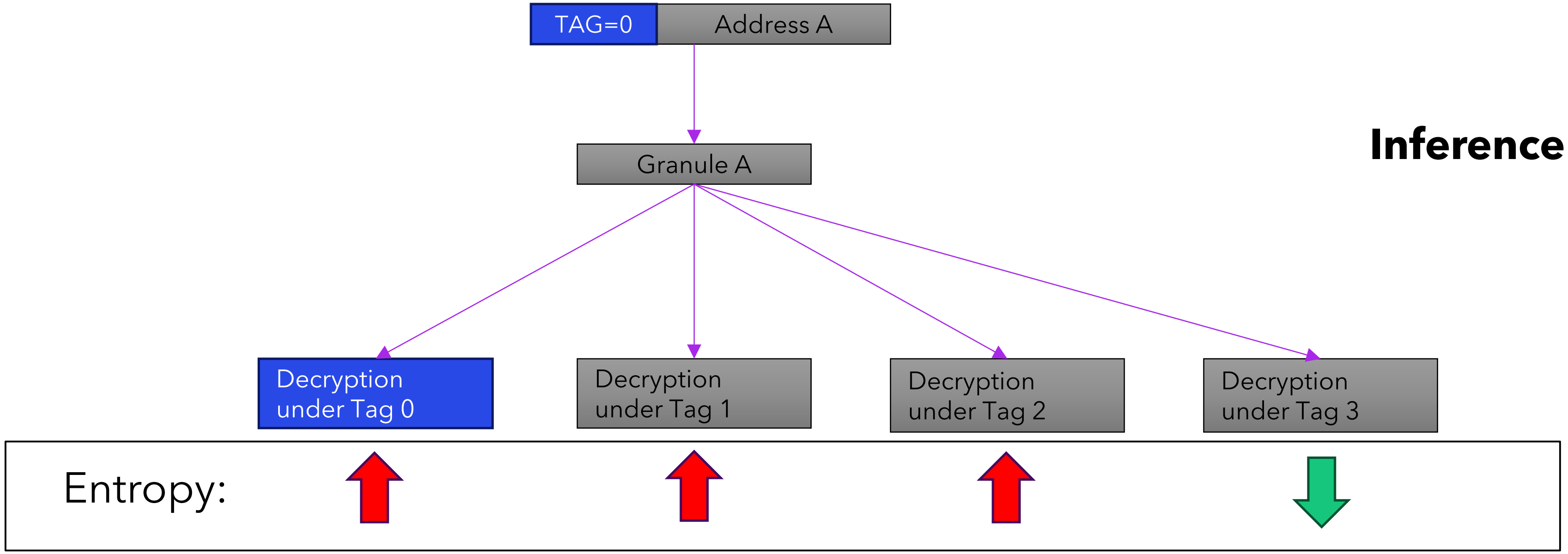
Data still protected by encryption.

Inference: Ambiguous

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Inferring Integrity

(Assume 2-bit tags)

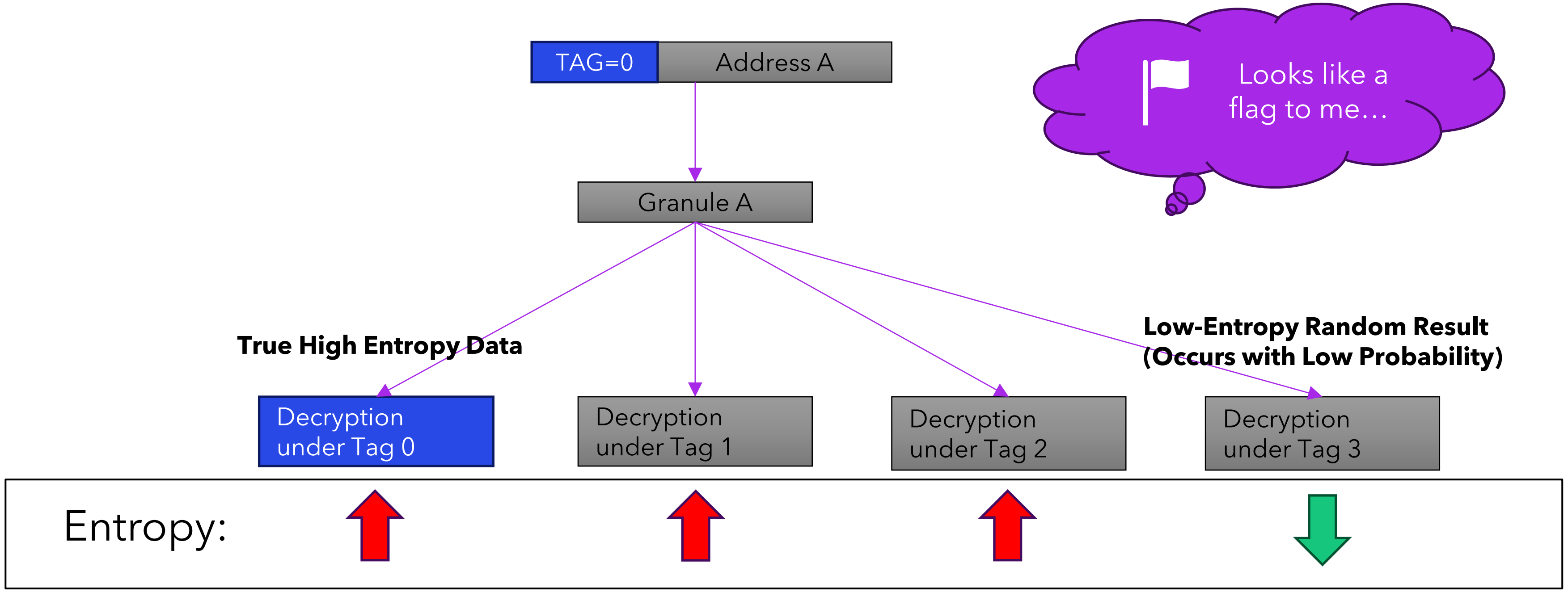


Inference: Potential Violation

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

False Positive Flagging

(Assume 2-bit tags)



This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

False Positive Table

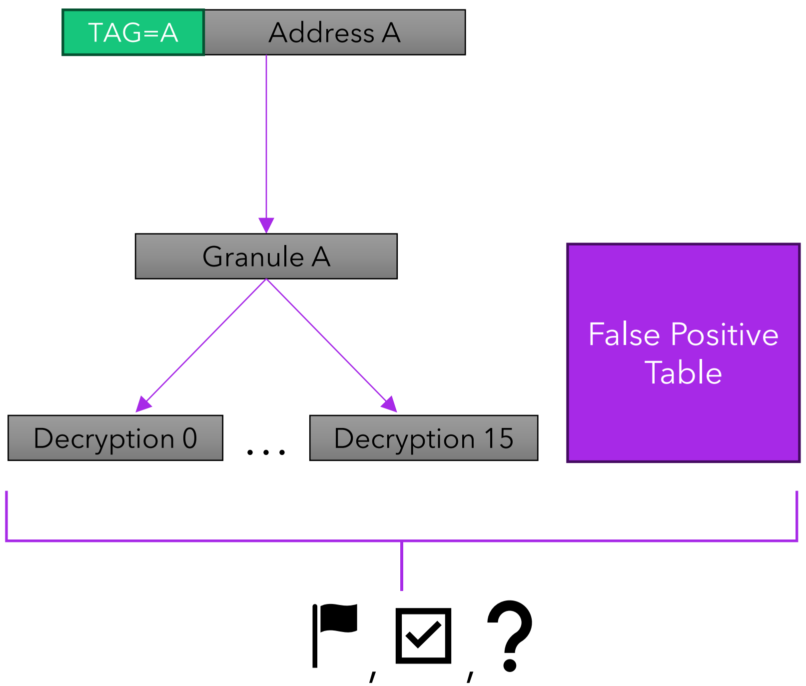
- The false positive table is the **only introduced state**.

- The table's expected size **linearly correlates** with the probability incorrect decryptions display low entropy, which can be tuned via entropy test parameters

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

CMT Execution Path

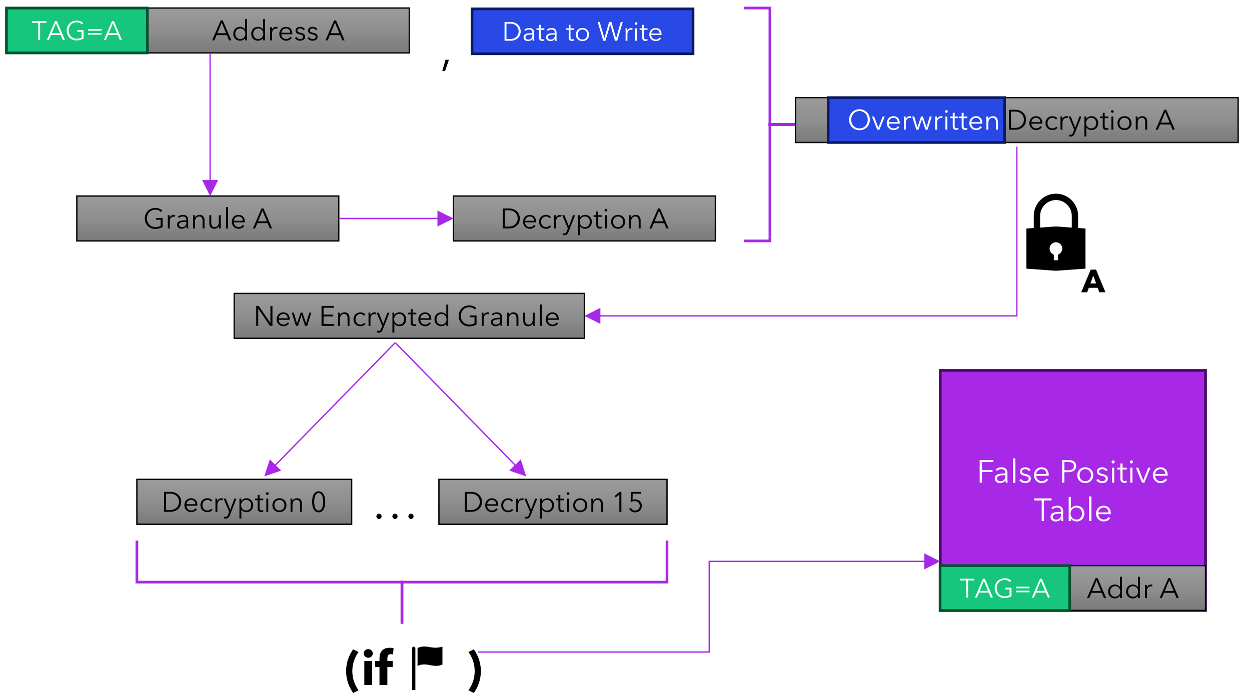
Step 1: Verify Access (Read and Writes)



This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

CMT Execution Path

Step 2: Catch False Positives (Write Only)



This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Efficacy Experiment

Aimed at quantifying **integrity coverage** and **lookup overhead** across active granules:

- Whenever an allocated granule is accessed by the workload, log or update its integrity outcome in a **memory map**.
- Each time a granule is freed, remove the granule from the map.
- **Every million accesses**, record the verification status of the entire active memory image.
- Calculate the **geometric mean** over these data points to report average portions for each workload.

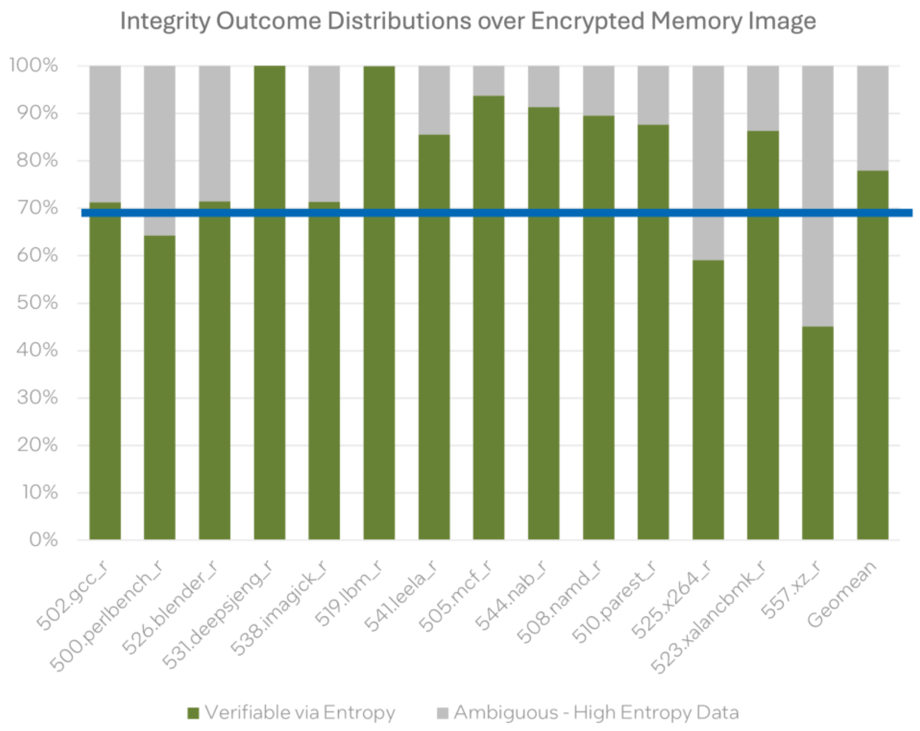
Implemented CMT using Intel Simics-based simulator

Conducted over SPEC CPU2017 testing suite

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Efficacy over Active Memory Image

- Geometric Mean of 80% coverage across workloads
- Less than <0.1% of operations resulted in a lookup
- Only 3 workloads fall below 70%, all of which store large amounts of high entropy data (ex. xz)



Conclusion

Cryptographic Memory Tagging achieves memory safety and data integrity with minimal overhead, providing a comprehensive, stateless approach to protecting modern computing systems.

Open Questions include:

- More nuanced, workload specific entropy tests
- Measuring entropy relative to other decryptions
- Hardware design brainstorming

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Cryptographic Memory Tagging

Towards Stateless Integrity

HASP ’24 | November 2, 2024

Presenter: Bharath Namboothiry (UPenn, Intel Labs)

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Appendix: P-Statelessness

- Let **P** be the probability that a uniformly random string (an incorrect decryption) registers as low entropy.
- In expectation, at most **P** portion of active memory granules have a lookup table entry
- We refer to such a scheme as **P-stateless**

- Goals choosing an entropy test and parameters:
  - Minimize **P**
  - Maintain ability to identify low entropy data

This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

Appendix: Byte Collision Test

Parameters: Granularity **g**, threshold **t**

- Input: a granule of length **g**
  - Initialize counter to 0
  - Iterate through bytes of granule. For each byte that is a repeat, increment counter by 1
  - Return **high entropy** if counter < t, and **low entropy** otherwise

  - We can express **P**, the probability of an incorrect low entropy result as:

$$P_{\mathrm{BYTE}}(g,t) = 1 - \left(1 - \frac{\sum_{n=1}^{g-t}\left[\binom{256}{n}\sum_{k=0}^{n}(-1)^k\binom{n}{k}(n-k)^g\right]}{256^g}\right)^{15}$$

This material is based u … 1-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.

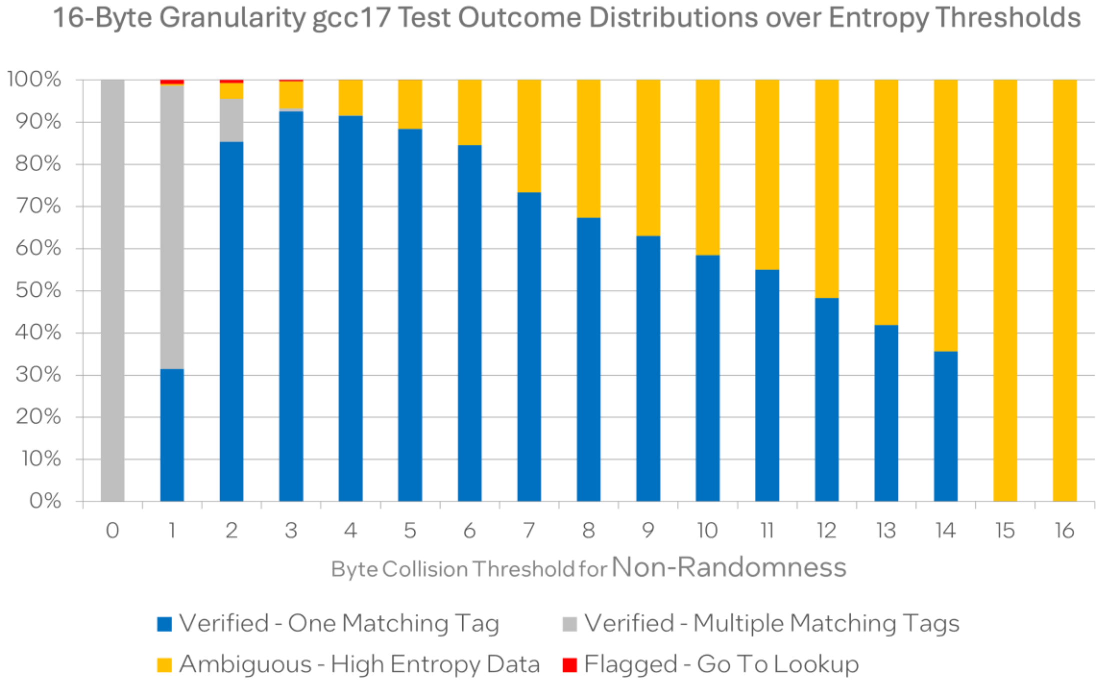
Appendix: Parameter Selection

Choice of **g = 16**:

- Generated stable entropy results
- Consistent with SOTA (ARM MTE)

Choice of **t = 4**:

- Reduces **P** while maintaining efficacy
- Larger **t** values can further minimize **P**, but trade off efficacy



This material is based upon work supported by the Naval Information Warfare Center Pacific and the Defense Advanced Research Project Agency under Prototype Other Transaction Agreement No. N66001-23-9-4004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Information Warfare Center Pacific or the Defense Advanced Research Project Agency.