RATS Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 7, 2020

H. Birkholz
Fraunhofer SIT
M. Wiseman
GE Global Research
H. Tschofenig
ARM Ltd.
N. Smith
Intel
M. Richardson
Sandelman Software Works
November 04, 2019

Remote Attestation Procedures Architecture
draft-birkholz-rats-architecture-03

Abstract

An entity (a relying party) requires a source of truth and evidence about a remote peer to assess the peer's trustworthiness. The evidence is typically a believable set of claims about its host, software or hardware platform. This document describes an architecture for such remote attestation procedures (RATS).

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Remote Attestation provides a way for an entity (the Relying Party) to determine the health and provenance of an endpoint/host (the Attester). Knowledge of the health of the endpoint allows for a determination of trustworthiness of the endpoint.

1.1. Motivation

The IETF has long spent it's time focusing on threats to the communication channel (see [RFC3552] and [DOLEV-YAO]), assuming that endpoints could be trusted and were under the observation of trusted, well-trained professionals. This assumption has not been true since the days of the campus mini-computer. For some time after the desktop PC became ubiquitous, the threat to the endpoints has been dealt with as an internal matter, with generally poor results. Enterprises have done some deployment of Network Endpoint Assessment ([RFC5209]) to assess the security posture about an endpoint, but it has not been ubiquitous.

The movement towards personal mobile devices ("smartphones") and the continuing threat from unmanaged residential desktops has resulted in a renewed interest in standardizing internet-scale endpoint remote attestation. Additionally, the rise of the Internet of Things (IoT) has made this issue even more critical: some skeptics have even renamed it to the Internet of Threats [iothreats] :-) IoT devices have poor or non-existent user interfaces, as such as there are not even good ways to assess the health of the devices manually: a need to determine the health via remote attestation is now critical.

In addition to the health of the device, knowledge of its provenance helps to determine the level of trust, and prevents attacks to the supply chain.

1.2. Opportunities

The Trusted Platform Module (TPM) is now a commonly available peripheral on many commodity compute platforms, both servers and desktops. Smartphones commonly have either an actual TPM, or have the ability to emulate one in software running in a Trusted Execution Environment [I-D.ietf-teep-architecture]. There are now few barriers to creating a standards-based system for remote attestation procedures.

A number of niche solutions have emerged that provide for use-case specific remote attestation, but none have the generality needed to be used across the Internet.

1.3. Overview of Document

The architecture described in this document (along with the accompanying solution and reference documents) enables the use of common formats for communicating Claims about an Attester to a Relying Party. [FIXME Attester? Flows? To what end?]

Existing transports were not designed to carry attestation Claims. It is therefore necessary to design serializations of Claims that fit into a variety of transports, for instance: X.509 certificates, TLS negotiations, YANG modules or EtherNet/IP. There are also new, greenfield uses for remote attestation. Transport and serialization of these can be done without retrofitting. This is (will be) described in [INSERT reference to adopted document on transport].

While it is not anticipated that the existing niche solutions described in the use cases section Section 3 will exchange claims directly, the use of a common format enables common code. As some of the code needs to be in intentionally hard to modify trusted modules, the use of a common formats and transfer protocols significantly reduces the cost of adoption to all parties. This commonality also significantly reduces the incidence of critical bugs.

In some environments the collection of Evidence by the Attester to be provided to the Verifier is part of an existing protocol: this document does not change that, rather embraces those legacy mechanisms as part of the specification. This is an evolutionary path forward, not revolutionary. Yet in other greenfield environments, there is a desire to have a standard for Evidence as well as for Attestation Results, and this architecture outlines how that is done.

This introduction gives an overview of the message flows and roles involved. Following this, is a terminology section that is referenced normatively by other documents and is a key part of this document. There is then a section on use cases and how they leverage the roles and workflows described.

In this document, terms defined within this document are consistently Capitalized [work in progress. please raise issues, if there are Blatant inconsistencies].

Current verticals that use remote attestation include:

- o The Trusted Computing Group "Network Device Attestation Workflow" [I-D.fedorkow-rats-network-device-attestation]

- o Android Keystore [keystore]

- o Fast Identity Online (FIDO) Alliance attestation [fido]

- o A number of Intel SGX niche systems based upon OTRP.

1.4. RATS in a Nutshell

1. Remote Attestation message flows typically convey Claims that contain the trustworthiness properties associated with an Attested Environment (Evidence).

2. A corresponding provisioning message flows conveys Reference trustworthiness claims that can be compared with attestation Evidence. Reference Values typically consist of firmware or software digests and details about what makes the attesting module a trusted source of Evidence.

3. Relying Parties are performing tasks such as managing a resource, controlling access, and/or managing risk. Attestation Results helps Relying Parties determine levels of trust.

1.5. Remote Attestation Workflow

The logical information flow is from Attester to Verifier to Relying Party. There are variations presented below on how this integrates into actual protocols.

```
                            ************
                            * Asserter *
                            ************
                                  |
                                  |
                                  |Known-Good-Values
                                  |Endorsements
                                  |
                                  v
                          .---------------.
                          |   Verifier    |
               .--------->|               |----------.
               |          |               |          |
               |          '_______________'          |
               |                                     |
               |Evidence                             |Attestation Results
               |                                     |(Claims)
               |                                     |
               |                                     v
       .---------------.                     .---------------.
       |   Attester    |                     | Relying Party |
       |               |                     |               |
       |               |                     |               |
       '_______________'                     '_______________'
```

Figure 1: RATS Workflow

In the architecture shown above, specific content items (payload conveyed in message flows) are identified:

- o Evidence is as set of believable Claims about distinguishable Environments made by an Attester.

- o Known-Good-Values are reference Claims used to appraise Evidence by an Verifier.

- o Endorsements are reference Claims about the type of protection that enables an Attester to create believable Evidence. Endorsements enable trust relationships towards system components or environments Evidence cannot be created for by an Attester.

- o Attestation Results are the output from the appraisal of Evidence, Known-Good-Values and Endorsements and are consumed by Relying Parties.

Attestation Results are the output of RATS.

Assessment of Attestation Results is be multi-faceted and out-of-scope for the architecture.

If appropriate Endorsements about the Attester are available, Known-Good-Values about the Attester are available, and if the Attester is capable of creating believable Evidence - then the Verifier is able to create Attestation Results that enable Relying Parties to establish a level of confidence in the trustworthiness of the Attester.

The Asserter role and the format for Known-Good-Values and Endorsements are not subject to standardization at this time. The current verticals already include provisions for encoding and/or distributing these objects.

1.6. Message Flows

Two distinct flows have been identified for passage of Evidence and production of Attestation Results. It is possible that there are additional situations which are not captured by these two flows.

1.6.1. Passport Model

In the Passport Model message flow the Attester provides it’s Evidence directly to the Verifier. The Verifier will evaluate the Evidence and then sign an Attestation Result. This Attestation Result is returned to the Attester, and it is up to the Attester to communicate the Attestation Result (potentially including the Evidence, if disclosable) to the Relying Party.

```
            ************
            * Asserter *
            ************
                  |Known-Good-Values
                  |Endorsements
                  |
                  v
          .---------------.
          |   Verifier    |
          |               |
          |               |
          '------------|--'
             ^         |
             |         |Attestation Results
     Evidence|         |(Claims)
             |         |
             |         |
             |         v
          .---|-----------.                     .---------------.
          |   Attester    |                     | Relying Party |
          |               ----------------------->              |
          |               | Attestation Results |               |
          '---------------' (Claims)            '---------------'
```

Figure 2: RATS Passport Flow

This flow is named in this way because of the resemblance of how Nations issue Passports to their citizens. The nature of the Evidence that an individual needs to provide to it's local authority is specific to the country involved. The citizen retains control of the resulting document and presents it to other entities when it needs to assert a citizenship or identity claim.

1.6.2. Background Check

In the Background-Check message flow the Attester provides it's Evidence to the Relying Party. The Relying Party sends this evidence to a Verifier of its choice. The Verifier will evaluate the Evidence and then sign an Attestation Result. This Attestation Result is returned to the Relying Party, which processes it directly.

```
                                            ************
                                            * Asserter *
                                            ************
                                                  |Known-Good-Values
                                                  |Endorsements
                                                  |
                                                  v
                                          .________________.
                                          |   Verifier    |
                                          |               |
                                          '__^_________|__'
                                             |         |
                                             |         |Attestation Results
                                    Evidence |         |(Claims)
                                             |         |
                                             |         v
    .________________.                    .__|____________.
    |   Attester    |      Evidence       | Relying Party |
    |               ---------------------->               |
    |               |                     |               |
    '_______________'                     '_______________'
```

Figure 3: RATS Background Check Flow

This flow is named in this way because of the resemblance of how employers and volunteer organizations perform background checks. When a prospective employee provides claims about education or previous experience, the employer will contact the respective institutions or former employers to validate the claim. Volunteer organizations often perform police background checks on volunteers in order to determine the volunteer's trustworthiness.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Appraisal: A Verifier process that compares Evidence to Reference values while taking into account Endorsements and produces Attestation Results.

Asserter: See Section 5.3.1.2.

Attester: See Section 5.3.1.1.

Attested Environment: A target environment that is observed or controlled by an Attesting Environment.

Attesting Environment: An environment capable of making trustworthiness Claims about an Attested Environment.

Background-Check Message Flow: An attestation workflow where the Attester provides Evidence to a Relying Party, who consults one or more Verifiers who supply Attestation Results to the Relying Party. See Section 1.6.2.

Claim: A statement about the construction, composition, validation or behavior of an Entity that affects trustworthiness. Evidence, Reference Values and Attestation Results are expressions that consists of one or more Claims.

Conveyance: The process of transferring Evidence, Reference Values and Attestation Results between Entities participating in attestation workflow.

Entity: A device, component (see System Component [RFC4949]), or environment that implements one or more Roles.

Evidence: See Section 5.3.2.1.

Passport Message Flow: An attestation workflow where the Attester provides Evidence to a Verifier who returns Attestation Results that are then forwarded to one or more Relying Parties. See Section 1.6.1.

Reference Values: See Section 5.3.2.2. Also referred to as Known-Good-Values.

Relying Party: See Section 5.3.1.4.

Attestation Results: See Section 5.3.2.3.

Role: A function or process in an attestation workflow, typically described by: Attester, Verifier, Relying Party and Asserter.

Verifier: See Section 5.3.1.3.

3. Reference use cases

This section provides an overview of a number of distinct use cases that benefit from a standardized claim format. In addition to outlining the user, the specific message flow is identified from among the flows detailed in Section 1.6.

3.1. Device Capabilities/Firmware Attestation

This is a large category of claims that includes a number of subcategories, not detailed here.

Use case name: Device Identity

Who will use it: Network Operators, larger enterprises

Attester: varies

Message Flow: sometimes passport and sometimes background check

Relying Party: varies

Description: Network operators want a trustworth report of identity and version of information of the hardware and software on the machines attached to their network. The process starts with some kind of Root of Trust that provides device identity and protected storage for measurements. The mechanism performs a series of measurements, and expresses this with an attestation as to the hardware and firmware/software which is running.

This is a general description for which there are many specific use cases, including [I-D.fedorkow-rats-network-device-attestation] section 1.2, "Software Inventory"

3.2. IETF TEEP WG Use-Case

Use case name: TAM validation

Who will use it: The TAM server

Message Flow: background check

Attester: Trusted Execution Environment (TEE)

Relying Party: end-application

Description: The "Trusted Application Manager (TAM)" server wants to verify the state of a TEE, or applications in the TEE, of a device. The TEE attests to the TAM, which can then decide whether to install sensitive data in the TEE, or whether the TEE is out of compliance and the TAM needs to install updated code in the TEE to bring it back into compliance with the TAM's policy.

3.3. Safety Critical Systems

Use case name: Safety Critical Systems

Who will use it: Power plants and other systems that need to assert their current state, but which can not accept any inputs from the outside. The corollary system is a black-box (such as in an aircraft), which needs to log the state of a system, but which can never initiate a handshake.

Message Flow: background check

Attester: web services and other sources of status/sensor information

Relying Party: open

Claims used as Evidence: the beginning and ending time as endorsed by a Time Stamp Authority, represented by a time stamp token. The real time clock of the system itself. A Root of Trust for time; the TPM has a relative time from startup.

Description: These requirements motivate the creation of the Time-Base Unidirectional Attestation (TUDA) [I-D.birkholz-rats-tuda], the output of TUDA is typically a secure audit log, where freshness is determined by synchronization to a trusted source of external time.

The freshness is preserved in the Evidence by the use of a Time Stamp Authority (TSA) which provides Time Stamp Tokens (TST).

3.4. Virtualized Multi-Tenant Hosts

Use case name: Multi-Tenant Hosts

Who will use it: Virtual machine systems

Message Flow: passport

Attester: virtual machine hypervisor

Relying Party: network operators

Description: The host system will do verification as per Section 3.1

The tenant virtual machines will do verification as per Section 3.1.

The network operator wants to know if the system _as a whole_ is free of malware, but the network operator is not allowed to know who the tenants are.

This is contrasted to the Chassis + Line Cards case (To Be Defined: TBD).

Multiple Line Cards, but a small attestation system on the main card can combine things together. This is a kind of proxy.

3.5. Cryptographic Key Attestation

Cryptographic Attestion includes subcategories such as Device Type Attestation (the FIDO use case), and Key storage Attestation (the Android Keystore use case), and End-User Authorization.

Use case name: Key Attestation

Who will use it: network authentication systems

Message Flow: passport

Attester: device platform

Relying Party: internet peers

Description: The relying party wants to know how secure a private key that identifies an entity is. Unlike the network attestation, the relying party is not part of the network infrastructure, nor do they necessarily have a business relationship (such as ownership) over the end device.

The Device Type Attestation is provided by a Firmware TPM performing the Verifier function, creating Attestation Results that indicate a particular model/type of device. In TCG terms, this is called Implicit Attestation, in this case the Attested Environment is the (smartphone) Rich Execution Environment (REE) ([I-D.ietf-teep-architecture] section 2), and the Attesting Environment is within the TEE.

3.6. Geographic Evidence

Use case name: Location Evidence

Who will use it: geo-fenced systems

Message Flow: passport (probably)

Attester: secure GPS system(s)

Relying Party: internet peers

Description: The relying party wants to know the physical location (on the planet earth, using a geodetic system) of the device. This may be provided directly by a GPS/GLONASS/BeiDou/Galileo module that is incorporated into a TPM. This may also be provided by collecting other proximity messages from other device that the relying party can form a trust relationship with.

3.7. Device Provenance Attestation

Use case name: RIV - Device Provenance

Who will use it: Industrial IoT devices

Message Flow: passport

Attester: network management station

Relying Party: a network entity

Description: A newly manufactured device needs to be onboarded into a network where many if not all device management duties are performed by the network owner. The device owner wants to verify the device originated from a legitimate vendor. A cryptographic device identity such as an IEEE802.1AR is embedded during manufacturing and a certificate identifying the device is delivered to the owner onboarding agent. The device authenticates using its 802.1AR IDevID to prove it originated from the expected vendor.

The device chain of custody from the original device manufacturer to the new owner may also be verified as part of device provenance attestation. The chain of custody history may be collected by a cloud service or similar capability that the supply chain and owner agree to use.

[I-D.fedorkow-rats-network-device-attestation] section 1.2 refers to this as "Provable Device Identity", and section 2.3 details the parties.

4. Conceptual Overview

In network protocol exchanges, it is often the case that one entity (a Relying Party) requires an assessment of the trustworthiness of a remote entity (an Attester or specific system components [RFC4949] thereof). Remote ATtestation procedureS (RATS) enable Relying Parties to establish a level of confidence in the trustworthiness of Attesters through the

- Creation,
- Conveyance, and
- Appraisal

of attestation Evidence.

Qualities of Evidence: Evidence is composed of Claims about trustworthiness (the set of Claims is unbounded). The system characteristics of Attesters - the Environments they are composed-of, and their continuous development - have an impact on the veracity of trustworthiness Claims included in valid Evidence.

Valid Evidence about the intactness of an Attester must be impossible to create by an untrustworthy or compromised Environment of an Attester.

Qualities of Environments: The resilience of Environments that are part of an Attester can vary widely - ranging from those highly resistant to attacks to those having little or no resistance to attacks. Configuration options, if set poorly, can result in a highly resistant environment being operationally less resistant. When a trustworthy Environment changes, it is possible that it transitions from being trustworthy to being untrustworthy.

An untrustworthy or compromised Environment must never be able to create valid Evidence expressing the intactness of an Attester.

The architecture provides a framework for anticipating when a relevant change with respect to a trustworthiness attribute occurs, what exactly changed and how relevant it is. The architecture also creates a context for enabling an appropriate response by applications, system software and protocol endpoints when changes to trustworthiness attributes do occur.

Detailed protocol specifications for message flows are defined in separate documents.

4.1. Two Types of Environments

An Attester produces Evidence about its own integrity, which means "it measures itself". To disambiguate this recursive or circular looking relationships, two types of Environments inside an Attester are distinguished:

The attest-ED Environments and the attest-ING Environments.

Attested Environments are measured. They provide the raw values and the information to be represented in Claims and ultimately expressed as Evidence.

Attesting Environments conduct the measuring. They collect the Claims, format them appropriately, and typically use key material and cryptographic functions, such as signing or cipher algorithms, to create Evidence.

Attesting Environments use system components that have to be trusted. As a result, Evidence includes Claims about the Attested and the Attesting Environments. Claims about the Attested Environments are appraised using Reference Values and Claims about the Attesting Environments are appraised using Endorsements. It is not mandated that both Environments have to be separate, but it is highly encouraged. Examples of separated Environments that can be used as Attesting Environments include: Trusted Execution Environments (TEE), embedded Secure Elements (eSE), or Hardware Security Modules (HSM).

In summary, the majority of the creation of evidence can take place in an Attested Environments. Exemplary duties include the collection and formatting of Claim values, or the trigger for creating Evidence. A trusted sub-set of the creation of evidence can take place in an Attesting Environment, that provide special protection with respect to key material, identity documents, or primitive functions to create the Evidence itself.

4.2. Evidence Creation Prerequisites

One or more Environments that are part of an Attester must be able to conduct the following duties in order to create Evidence:

- o monitoring trustworthiness attributes of other Environments,

- o collecting trustworthiness attributes and create Claims about them,

- o serialize Claims using interoperable representations,

- o provide integrity protection for the sets of Claims, and

- o add appropriate attestation provenance attributes about the sets of Claims.

4.3. Trustworthiness

The trustworthiness of an Attester and therefore the believability of the Evidence it creates relies on the protection methods in place to shield and restrict the use of key material and the duties conducted by the Attesting Environment. In order to assess trustworthiness effectively, it is mandatory to understand the trustworthiness properties of the environments of an Attester. The corresponding appraisal of Evidence that leads to such an assessment of trustworthiness is the duty of a Verifier.

Trusting the assessment of a Verifier might com frome trusting the Verifier's key material (direct trust), or trusting an Entity that the Verifier is associated with via a certification path (indirect trust).

The trustworthiness of corresponding Attestation Results also relies on trust towards manufacturers and those manufacturer's hardware in order to assess the integrity and resilience of that manufacturer's devices.

A stronger level of security comes when information can be vouched for by hardware or by (unchangeable) firmware, especially if such hardware is physically resistant to hardware tampering. The component that is implicitly trusted is often referred to as a Root of Trust.

4.4. Workflow

The basic function of RATS is creation, conveyance and appraisal of attestation Evidence. An Attester creates attestation Evidence that are conveyed to a Verifier for appraisal. The appraisals compare Evidence with expected Known-Good-Values obtained from Asserters (e.g. Principals that are Supply Chain Entities). There can be multiple forms of appraisal (e.g., software integrity verification, device composition and configuration verification, device identity and provenance verification). Attestation Results are the output of appraisals. Attestation Results are signed and conveyed to Relying Parties. Attestation Results provide the basis by which the Relying Party may determine a level of confidence to place in the application data or operations that follow.

The architecture defines attestation Roles: Attester, Verifier, Asserter and Relying Party. Roles exchange messages, but their structure is not defined in this document. The detailed definition of the messages is in an appropriate document, such as [I-D.ietf-rats-eat] or other protocols to be defined. Roles can be combined in various ways into Principals, depending upon the needs of the use case. Information Model representations are realized as data structure and conveyance protocol specifications.

4.5. Interoperability between RATS

The RATS architecture anticipates use of information modeling techniques that describe computing structures - their components/computational elements and corresponding capabilities - so that verification operations may rely on the information model as an interoperable way to navigate the structural complexity.

5. RATS Architecture

5.1. Goals

RATS architecture has the following goals:

- Enable semantic interoperability of attestation semantics through information models about computing environments and trustworthiness.

- Enable data structure interoperability related to claims, endpoint composition / structure, and end-to-end integrity and confidentiality protection mechanisms.

- Enable programmatic assessment of trustworthiness. (Note: Mechanisms that manage risk, justify a level of confidence, or determine a consequence of an attestation result are out of scope).

- Provide the building blocks, including Roles and Principals that enable the composition of service-chains/hierarchies and workflows that can create and appraise evidence about the trustworthiness of devices and services.

- Use-case driven architecture and design (see [I-D.richardson-rats-usecases] and Section 3)

- Terminology conventions that are consistently applied across RATS specifications.

- Reinforce trusted computing principles that include attestation.

5.2. Attestation Principles

Specifications developed by the RATS working group apply the following principles:

- o Freshness - replay of previously asserted Claims about an Attested Environment can be detected.

- o Identity - the Attesting Environment is identifiable (non-anonymous).

- o Context - the Attested Environment is well-defined (unambiguous).

- o Provenance - the origin of Claims with respect to the Attested and Attesting Environments are known.

- o Validity - the expected lifetime of Claims about an Attested Environment is known.

- o Veracity - the believability (level of confidence) of Claims is based on verifiable proofs.

5.3. Attestation Workflow

Attestation workflow helps a Relying Party make better decisions by providing insight about the trustworthiness of endpoints participating in a distributed system. The workflow consists primarily of four roles; Relying Party, Verifier, Attester and Asserter. Attestation messages contain information useful for appraising the trustworthiness of an Attester endpoint and informing the Relying Party of the appraisal result.

This section details the primary roles of an attestation workflow and the messages they exchange.

5.3.1. Roles

An endpoint system (a.k.a., Entity) may implement one or more attestation Roles to accommodate a variety of possible message flows. Exemplary message flows are described in Section 1.6.1 and Section 1.6.2. Role messages are secured by the Entity that generated it. Entities possess credentials (e.g., cryptographic keys) that authenticate, integrity protect and optionally confidentiality protect attestation messages.

5.3.1.1. Attester

The Attester consists of both an Attesting Environment and an Attested Environment. In some implementations these environments may be combined. Other implementations may have multiples of Attesting and Attested environments. Although endpoint environments can be complex, and that complexity is security relevant, the basic function of an Attester is to create Evidence that captures operational conditions affecting trustworthiness.

5.3.1.2. Asserter

The Asserter role is out of scope. The mechanism by which an Asserter communicates Known-Good-Values to a Verifier is also not subject to standardization. Users of the RATS architecture are assumed to have pre-existing mechanisms.

5.3.1.3. Verifier

The Verifier workflow function accepts Evidence from an Attester and accepts Reference Values from one or more Asserters. Reference values may be supplied a priori, cached or used to created policies. The Verifier performs an appraisal by matching Claims found in Evidence with Claims found in Reference Values and policies. If an attested Claim value differs from an expected Claim value, the Verifier flags this as a change possibly impacting trust level.

Endorsements may not have corresponding Claims in Evidence (because of their intrinsic nature). Consequently, the Verifier need only authenticate the endpoint and verify the Endorsements match the endpoint identity.

The result of appraisals and Endorsements, informed by owner policies, produces a new set of Claims that a Relying Party is suited to consume.

5.3.1.4. Relying Party

A Role in an attestation workflow that accepts Attestation Results from a Verifier that may be used by the Relying Party to inform application specific decision making. How Attestation Results are used to inform decision making is out-of-scope for this architecture.

5.3.2. Role Messages

5.3.2.1. Evidence

Claims that are formatted and protected by an Attester.

Evidence SHOULD satisfy Verifier expectations for freshness, identity, context, provenance, validity, and veracity.

5.3.2.2. Reference Values

Reference-values are Claims that a manufacturer, vendor or other supply chain entity makes that affects the trustworthiness of an Attester endpoint.

Claims may be persistent properties of the endpoint due to the physical nature of how it was manufactured or may reflect the processes that were followed as part of moving the endpoint through a supply-chain; e.g., validation or compliance testing. This class of Reference-values is known as Endorsements.

Another class of Reference-values identifies the firmware and software that could be installed in the endpoint after its manufacture. A digest of the the firmware or software can be an effective identifier for keeping track of the images produced by vendors and installed on an endpoint. This class of Reference-value is referred to as Known-Good-Value (KGV).

Known-Good-Values: Claims about the Attested Environment. Typically, Known-Good-Value (KGV) Claims are message digests of firmware, software or configuration data supplied by various vendors. If an Attesting Environment implements cryptography, they include Claims about key material.

Like Claims, Known-Good-Values SHOULD satisfy a Verifier's expectations for freshness, identity, context, provenance, validity, relevance and veracity. Known-Good-Values are reference Claims that are - like Evidence - well formatted and protected (e.g. signed).

Endorsements: Claims about immutable and implicit characteristics of the Attesting Environment. Typically, endorsement Claims are created by manufacturing or supply chain entities.

Endorsements are intended to increase the level of confidence with respect to Evidence created by an Attester.

5.3.2.3. Attestation Results

Statements about the output of an appraisal of Evidence that are created, formatted and protected by a Verifier.

Attestation Results provide the basis for a Relying Party to establish a level of confidence in the trustworthiness of an Attester. Attestation Results SHOULD satisfy Relying Party expectations for freshness, identity, context, provenance, validity, relevance and veracity.

5.4. Principals (Entities?) - Containers for the Roles

[The authors are unhappy with the term Principal, and have been looking for something else. JOSE/JWT uses the term Principal]

Principals are Containers for the Roles.

Principals are users, organizations, devices and computing environments (e.g., devices, platforms, services, peripherals).

Principals may implement one or more Roles. Massage flows occurring within the same Principal are out-of-scope.

The methods whereby Principals may be identified, discovered, authenticated, connected and trusted, though important, are out-of-scope.

Principal operations that apply resiliency, scaling, load balancing or replication are generally believed to be out-of-scope.

```
+------------------+   +------------------+
|  Principal 1     |   |  Principal 2     |
|  +------------+  |   |  +------------+  |
|  |            |  |   |  |            |  |
|  |    Role A  |<-|---|->|    Role D  |  |
|  |            |  |   |  |            |  |
|  +------------+  |   |  +------------+  |
|                  |   |                  |
|  +-----+------+  |   |  +-----+------+  |
|  |            |  |   |  |            |  |
|  |    Role B  |<-|---|->|    Role E  |  |
|  |            |  |   |  |            |  |
|  +------------+  |   |  +------------+  |
|                  |   |                  |
+------------------+   +------------------+
```

Figure 4: Principals-Role Composition

Principals have the following properties:

- o Multiplicity - Multiple instances of Principals that possess the same Roles can exist.

- o Composition - Principals possessing different Roles can be combined into a singleton Principal possessing the union of Roles. Message flows between combined Principals is uninteresting.

- o Decomposition - A singleton Principal possessing multiple Roles can be divided into multiple Principals.

6. Privacy Considerations

The conveyance of Evidence and the resulting Attestation Results reveal a great deal of information about the internal state of a device. In many cases the whole point of the Attestation process is to provided reliable evidence about the type of the device and the firmware that the device is running. This information is particularly interesting to many attackers: knowing that a device is running a weak version of a the firmware provides a way to aim attacks better.

Just knowing the existence of a device is itself a disclosure.

Conveyance protocols must detail what kinds of information is disclosed, and to whom it is exposed.

7. Security Considerations

Evidence, Verifiable Assertions and Attestation Results SHOULD use formats that support end-to-end integrity protection and MAY support end-to-end confidentiality protection.

Replay attacks are a concern that protocol implementations MUST deal with. This is typically done via a Nonce Claim, but the details belong to the protocol.

All other attacks involving RATS structures are not explicitly addressed by the architecture.

Additional security protections MAY be required of conveyance mechanisms. For example, additional means of authentication, confidentiality, integrity, replay, denial of service and privacy protection of RATS payloads and Principals may be needed.

8. Acknowledgements

Dave Thaler created the concepts of "Passport" and "Background Check".

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

9.2. Informative References

[ABLP] Abadi, M., Burrows, M., Lampson, B., and G. Plotkin, "A Calculus for Access Control in Distributed Systems", Springer Annual International Cryptology Conference, page 1-23, DOI 10.1.1.36.691, 1991.

[DOLEV-YAO] Dolev, D. and A. Yao, "On the security of public key protocols", IEEE Transactions on Information Theory Vol. 29, pp. 198-208, DOI 10.1109/tit.1983.1056650, March 1983.

[fido] FIDO Alliance, ., "FIDO Specification Overview", 2019, <https://fidoalliance.org/specifications/>.

[I-D.birkholz-rats-tuda] Fuchs, A., Birkholz, H., McDonald, I., and C. Bormann, "Time-Based Uni-Directional Attestation", draft-birkholz-rats-tuda-01 (work in progress), September 2019.

[I-D.fedorkow-rats-network-device-attestation] Fedorkow, G. and J. Fitzgerald-McKay, "Network Device Attestation Workflow", draft-fedorkow-rats-network-device-attestation-00 (work in progress), July 2019.

[I-D.ietf-rats-eat] Mandyam, G., Lundblade, L., Ballesteros, M., and J. O'Donoghue, "The Entity Attestation Token (EAT)", draft-ietf-rats-eat-01 (work in progress), July 2019.

[I-D.ietf-teep-architecture] Pei, M., Tschofenig, H., Wheeler, D., Atyeo, A., and D. Liu, "Trusted Execution Environment Provisioning (TEEP) Architecture", draft-ietf-teep-architecture-03 (work in progress), July 2019.

[I-D.richardson-rats-usecases]
Richardson, M., Wallace, C., and W. Pan, "Use cases for Remote Attestation common encodings", draft-richardson-rats-usecases-05 (work in progress), October 2019.

[iothreats]
GDN, ., "The Internet of Things or the Internet of threats?", 2016, <https://gcn.com/articles/2016/05/03/internet-of-threats.aspx>.

[keystore]
Google, ., "Android Keystore System", 2019, <https://developer.android.com/training/articles/keystore>.

[Lampson2007]
Lampson, B., "Practical Principles for Computer Security", IOSPress Proceedings of Software System Reliability and Security, page 151-195, DOI 10.1.1.63.5360, 2007.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <https://www.rfc-editor.org/info/rfc3552>.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", FYI 36, RFC 4949, DOI 10.17487/RFC4949, August 2007, <https://www.rfc-editor.org/info/rfc4949>.

[RFC5209] Sangster, P., Khosravi, H., Mani, M., Narayan, K., and J. Tardo, "Network Endpoint Assessment (NEA): Overview and Requirements", RFC 5209, DOI 10.17487/RFC5209, June 2008, <https://www.rfc-editor.org/info/rfc5209>.

Authors' Addresses

Henk Birkholz
Fraunhofer SIT
Rheinstrasse 75
Darmstadt 64295
Germany

Email: henk.birkholz@sit.fraunhofer.de

Monty Wiseman
GE Global Research
USA

Email: monty.wiseman@ge.com

Hannes Tschofenig
ARM Ltd.
110 Fulbourn Rd
Cambridge CB1 9NJ
UK

Email: hannes.tschofenig@gmx.net

Ned Smith
Intel Corporation
USA

Email: ned.smith@intel.com

Michael Richardson
Sandelman Software Works
Canada

Email: mcr+ietf@sandelman.ca

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 20, 2018

G. Mandyam
L. Lundblade
M. Ballesteros
J. O'Donoghue
Qualcomm Technologies Inc.

May 19, 2018

The Entity Attestation Token (EAT)
draft-mandyam-eat-00

Abstract

An attestation format based on concise binary object representation (CBOR) is proposed that is suitable for inclusion in a CBOR Web Token (CWT), know as the Entity Attestation Token (EAT). The associated data can be used by a relying party to assess the security state of a remote device or module.

Contributing

TBD

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Remote device attestation is fundamental service that allows a remote device such as a mobile phone, an Internet-of-Things (IoT) device, or other endpoint to prove itself to a relying party, a server or a service. This allows the relying party to know some characteristics about the device and decide whether it trusts the device.

Remote attestation is a fundamental service that can underly other protocols and services that need to know about the trustworthiness of the device before proceeding. One good example is biometric authentication where the biometric matching is done on the device. The relying party needs to know that the device is one that is known to do biometric matching correctly. Another example is content protection where the relying party wants to know the device will protect the data. This generalizes on to corporate enterprises that might want to know that a device is trustworthy before allowing corporate data to be accessed by it.

The notion of attestation here is large and may include, but is not limited to the following:

- Proof of the make and model of the device hardware (HW)
- Proof of the make and model of the device processor, particularly for security oriented chips
- Measurement of the software (SW) running on the device
- Configuration and state of the device
- Environmental characteristics of the device such as its GPS location

The required data format should be general purpose and extensible so that it can work across many use cases. This is why CBOR (see [RFC7049]) is choosen as the format -- it already supports a rich set of data types, and is both expressive and extensible. It translates well to JSON for good interoperation with web technology. It is compact and can work on very small IoT device. The format proposed here is small enough that a limited version can be implemented in pure hardware gates with no software at all. Moreover, the attestation data is defined in the form of claims that is the same as CBOR Web Token (CWT, see [RFC8392]). This is the motivation for defining the Entity Attestation Token, i.e. EAT.

1.1. Entity Overview

An "entity" can be any device or device subassembly ("submodule") that can generate its own attestation in the form of an EAT. The attestation should be cryptographically verifiable by the EAT consumer. An EAT at the device-level can be composed of several submodule EAT’s. It is assumed that any entity that can create an EAT does so by means of a dedicated root-of-trust (RoT).

Modern devices such as a mobile phone have many different execution environments operating with different security levels. For example it is common for a mobile phone to have an "apps" environment that runs an operating system (OS) that hosts a plethora of downloadable apps. It may also have a TEE (Trusted Execution Environment) that is distinct, isolated, and hosts security-oriented functionality like biometric authentication. Additionally it may have an eSE (embedded Secure Element) - a high security chip with defenses against HW attacks that can serve as a RoT. This device attestation format allows the attested data to be tagged at a security level from which it originates. In general, any discrete execution environment that has an identifiable security level can be considered an entity.

1.2. Use of CBOR and COSE

Fundamentally this attestation format is a verifiable data format. It is a collection of data items that can be signed by an attestation key, hashed, and/or encrypted. As per Section 7 of [RFC8392], the verification method is in the CWT using the CBOR Object Signing and Encryption (COSE) methodology (see [RFC8152]).

In addition, the reported attestation data could be determined within the secure operating environment or written to it from an external and presumably less trusted entity on the device. In either case, the source of the reported data must be identifiable by the relying party.

This attestation format is a single relatively simple signed message. It is designed to be incorporated into many other protocols and many other transports. It is also designed such that other SW and apps can add their own data to the message such that it is also attested.

1.3. EAT Operating Models

At least the following three participants exist in all EAT operating models. Some operating models have additional participants.

The Entity. This is the phone, the IoT device, the sensor, the sub-assembly or such that the attestation provides information about.

The Manufacturer. The company that made the entity. This may be a chip vendor, a circuit board module vendor or a vendor of finished consumer products.

The Relying Party. The server, service or company that makes use of the information in the EAT about the entity.

In all operating models, the manufacturer provisions some secret attestation key material (AKM) into the entity during manufacturing. This might be during the manufacturer of a chip at a fabrication facility (fab) or during final assembly of a consumer product or any time in between. This attestation key material is used for signing EATs.

In all operating models, hardware and/or software on the entity create an EAT of the format described in this document. The EAT is always signed by the attestation key material provisioned by the manufacturer.

In all operating models, the relying party must end up knowing that the signature on the EAT is valid and consistent with data from claims in the EAT. This can happen in many different ways. Here are some examples.

- o The EAT is transmitted to the relying party. The relying party gets corresponding key material (e.g. a root certificate) from the manufacturer. The relying party performs the verification.

- o The EAT is transmitted to the relying party. The relying party transmits the EAT to a verification service offered by the manufacturer. The server returns the validated claims.

- o The EAT is transmitted directly to a verification service, perhaps operated by the manufacturer or perhaps by another party. It verifies the EAT and makes the validated claims available to the relying party. It may even modify the claims in some way and re-sign the EAT (with a different signing key).

This standard supports all these operating models and does not prefer one over the other. It is important to support this variety of operating models to generaly facilitate deployment and to allow for some special scenarios. One special scenario has a validation service that is monetized, most likely by the manufacturer. In another, a privacy proxy service processes the EAT before it is transmitted to the relying party. In yet another, symmetric key material is used for signing. In this case the manufacturer should perform the verification, because any release of the key material would enable a participant other than the entity to create valid signed EATs.

1.4. What is Not Standardized

1.4.1. Transmission Protocol

EATs may be transmitted by any protocol. For example, they might be added in extension fields of other protocols, bundled into an HTTP header, or just transmitted as files. This flexibility is intentional to allow broader adoption. This flexibility is possible because EAT’s are self-secured with signing (and possibly additionally with encryption and anti-replay). The transmission protocol is not required to fulfill any additional security requirements.

For certain devices, a direct connection may not exist between the EAT-producing device and the Relying Party. In such cases, the EAT should be protected against malicious access. The use of COSE allows for signing and encryption of the EAT. Therefore even if the EAT is conveyed through intermediaries between the device and Relying Party, such intermediaries cannot easily modify the EAT payload or alter the signature.

1.4.2. Signing Scheme

The term "signing scheme" is used to refer to the system that includes end-end process of establishing signing attesation key material in the entity, signing the EAT, and verifying it. This might involve key IDs and X.509 certificate chains or something similar but different. The term "signing algorithm" refers just to the algorithm ID in the COSE signing structure. No particular signing algorithm or signing scheme is required by this standard.

There are three main implementation issues driving this. First, secure non-volatile storage space in the entity for the attestation key material may be highly limited, perhaps to only a few hundred bits, on some small IoT chips. Second, the factory cost of provisioning key material in each chip or device may be high, with even millisecond delays adding to the cost of a chip. Third, privacy-preserving signing schemes like ECDAA (Elliptic Curve Direct Anonymous Attestation) are complex and not suitable for all use cases.

Eventually some form of standardization of the signing scheme may be required. This might come in the form of another standard that adds to this document, or when there is clear convergence on a small number of signing schemes this standard can be updated.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

This document reuses terminology from JWT [RFC7519], COSE [RFC8152], and CWT [RFC8392].

StringOrURI. The "StringOrURI" term in this specification has the same meaning and processing rules as the JWT "StringOrURI" term defined in Section 2 of [RFC7519], except that it is represented as a CBOR text string instead of a JSON text string.

NumericDate. The "NumericDate" term in this specification has the same meaning and processing rules as the JWT "NumericDate" term defined in Section 2 of [RFC7519], except that it is represented as a CBOR numeric date (from Section 2.4.1 of [RFC7049]) instead of a JSON number. The encoding is modified so that the leading tag 1 (epoch-based date/time) MUST be omitted.

Claim Name. The human-readable name used to identify a claim.

Claim Key. The CBOR map key used to identify a claim.

Claim Value. The CBOR map value representing the value of the claim.

CWT Claims Set. The CBOR map that contains the claims conveyed by the CWT.

FloatOrNumber. The "FloatOrNumber" term in this specification is the type of a claim that is either a CBOR postivie integer, negative integer or floating point number.

Attestation Key Material (AKM). The key material used to sign the EAT token. If it is done symmetrically with HMAC, then this is a simple symmetric key. If it is done with ECC, such as an IEEE DevID [IDevID], then this is the private part of the EC key pair. If ECDAA is used, (e.g., as used by Ehnaced Privacy ID, i.e. EPID) then it is the key material needed for ECDAA.

3. The Claims

3.1. Universal Entity ID (UEID) Claim

UEID’s identify individual manufactured entities / devices such as a mobile phone, a water meter, a Bluetooth speaker or a networked security camera. It may identify the entire device or a submodule or subsystem. It does not identify types, models or classes of devices. It is akin to a serial number, though it does not have to be sequential.

It is identified by Claim Key X (X is TBD).

UEID’s must be universally and globally unique across manufacturers and countries. UEIDs must also be unique across protocols and systems, as tokens are intended to be embedded in many different protocols and systems. No two products anywhere, even in completely different industries made by two different manufacturers in two different countries. should have the same UEID (if they are not global and universal in this way then relying parties receiving them will have to track other characteristics of the device to keep devices distinct between manufacturers).

The UEID should be permanent. It should never change for a given device / entity. In addition, it should not be reprogrammable.

UEID’s are binary byte-strings (resulting in a smaller size than text strings). When handled in text-based protocols, they should be base-64 encoded.

UEID’s are variable length with a maximum size of 33 bytes (1 type byte and 256 bits). A receivers of a token with UEIDs may reject the token if a UEID is larger than 33 bytes.

UEID’s are not designed for direct use by humans (e.g., printing on the case of a device), so no textual representation is defined.

A UEID is a byte string. From the consumer’s view (the rely party) it is opaque wtih no bytes having any special meaning.

When the entity constructs the UEID, the first byte is a type and the following bytes the ID for that type. Several types are allowed to accommodate different industries and different manufacturing processes and to give options to avoid paying fees for certain types of manufacturer registrations.

Type Byte	Type Name	Specification
0x01	GUID	This is a 128 to 256 bit random number generated once and stored in the device. The GUID may be constructed from various identifiers on the device using a hash function or it may be just the raw random number. In any case, the random number must have entropy of at least 128 bits as this is what gives the global
0x02	IEEE EUI	This makes use of the IEEE company identification registry. An EUI is made up of an OUI and OUI-36 or a CID, different registered company identifiers, and some unique per-device identifier. EUIs are often the same as or similar to MAC addresses. (Note that while devices with multiple network interfaces may have multiple MAC addresses, there is only one UEID for a device) TODO: normative references to IEEE.
0x03	IMEI	TODO: figure how to specify IMEIs

Table 1: UEID Composition Types

The consumer (the Relying Party) of a UEID should treat a UEID as a completely opaque string of bytes and not make any use of its internal structure. For example they should not use the OUI part of a type 0x02 UEID to identify the manufacturer of the device. Instead they should use the OUI claim that is defined elsewhere. The reasons for this are:

- o UEIDs types may vary freely from one manufacturer to the next.
- o New types of UEIDs may be created. For example a type 0x04 UEID may be created based on some other manufacturer registration scheme.
- o Device manufacturers are allowed to change from one type of UEID to another anytime they want. For example they may find they can optimize their manufacturing by switching from type 0x01 to type

0x02 or vice versa. The main requirement on the manufacturer is that UEIDs be universally unique.

3.2. Origination (origination) Claims

This claim describes the parts of the device or entity that are creating the EAT. Often it will be tied back to the device or chip manufacturer. The following table gives some examples:

Name	Description
Acme-TEE	The EATs are generated in the TEE authored and configured by "Acme"
Acme-TPM	The EATs are generated in a TPM manufactured by "Acme"
Acme-Linux-Kernel	The EATs are generated in a Linux kernel configured and shipped by "Acme"
Acme-TA	The EATs are generated in a Trusted Application (TA) authored by "Acme"

The claim is represented by Claim Key X+15. It is type StringOrURI.

TODO: consider a more structure approach where the name and the URI and other are in separate fields.

TODO: This needs refinement. It is somewhat parallel to issuer claim in CWT in that it describes the authority that created the token.

3.3. OEM identification by IEEE OUI

This claim identifies a device OEM by the IEEE OUI. Reference TBD. It is a byte string representing the OUI in binary form in network byte order (TODO: confirm details).

Companies that have more than one IEEE OUI registered with IEEE should pick one and prefer that for all their devices.

Note that the OUI is in common use as a part of MAC Address. This claim is only the first bits of the MAC address that identify the manufacturer. The IEEE maintains a registry for these in which many companies participate. This claim is represented by Claim Key TBD.

3.4. Security Level (seclevel) Claim

EATs have a claim that roughly characterizes the device / entities ability to defend against attacks aimed at capturing the signing key, forging claims and at forging EATs. This is done by roughly defining four security levels as described below. This is similar to the security levels defined in the Metadata Service definied by the Fast Identity Online (FIDO) Alliance (TODO: reference).

These claims describe security environment and countermeasures available on the end-entity / client device where the attestation key reside and the claims originate.

This claim is identified by Claim Key X+2. The value is an integer between 1 and 4 as defined below.

1 - Unrestricted There is some expectation that implementor will protect the attestation signing keys at this level. Otherwise the EAT provides no meaningful security assurances.

2- Restricted Entities at this level should not be general-purpose operating environments that host features such as app download systems, web browsers and complex productivity applications. It is akin to the Secure Restricted level (see below) without the security orientation. Examples include a WiFi subsystem, an IoT camera, or sensor device.

3 - Secure Restriced Entities at this level must meet the critera defined by FIDO Allowed Restricted Operating Environments (TODO: reference). Examples include TEE’s and schemes using virtualization-based security. Like the FIDO security goal, security at this level is aimed at defending well against large-scale network / remote attacks against the device.

4 - Hardware Entities at this level must include substantial defense against physical or electrical attacks against the device itself. It is assumed any potential attacker has captured the device and can disassemble it. Example include TPMs and Secure Elements.

This claim is not intended as a replacement for a proper end-device security certification schemes such as those based on FIPS (TODO: reference) or those based on Common Criterion (TODO: reference). The claim made here is solely a self-claim made by the Entity Originator.

3.5. Nonce (nonce) Claim

The "nonce" (Nonce) claim represents a random value that can be used to avoid replay attacks. This would be ideally generated by the CWT consumer. This value is intended to be a CWT companion claim to the existing JWT claim **_IANAJWT_ (TODO: fix this reference). The nonce claim is identified by Claim Key X+3.

3.6. Secure Boot and Debug Enable State Claims

3.6.1. Secure Boot Enabled (secbootenabled) Claim

The "secbootenabled" (Secure Boot Enabled) claim represents a boolean value that indicates whether secure boot is enabled either for an entire device or an individual submodule. If it appears at the device level, then this means that secure boot is enabled for all submodules. Secure boot enablement allows a secure boot loader to authenticate software running either in a device or a submodule prior allowing execution. This claim is identified by Claim Key X+4.

3.6.2. Debug Disabled (debugdisabled) Claim

The "debugdisabled" (Debug Disabled) claim represents a boolean value that indicates whether debug capabilities are disabled for an entity (i.e. value of ’true’). Debug disablement is considered a prerequisite before an entity is considered operational. This claim is identified by Claim Key X+5.

3.6.3. Debug Disabled Since Boot (debugdisabledsincebboot) Claim

The "debugdisabledsinceboot" (Debug Disabled Since Boot) claim represents a boolean value that indicates whether debug capabilities for the entity were not disabled in any way since boot (i.e. value of ’true’). This claim is identified by Claim Key X+6.

3.6.4. Debug Permanent Disable (debugpermanentdisable) Claim

The "debugpermanentdisable" (Debug Permanent Disable) claim represents a boolean value that indicates whether debug capabilities for the entity are permanently disabled (i.e. value of ’true’). This value can be set to ’true’ also if only the manufacturer is allowed to enabled debug, but the end user is not. This claim is identified by Claim Key X+7.

3.6.5. Debug Full Permanent Disable (debugfullpermanentdisable) Claim

The "debugfullpermanentdisable" (Debug Full Permanent Disable) claim represents a boolean value that indicates whether debug capabilities for the entity are permanently disabled (i.e. value of 'true'). This value can only be set to 'true' if no party can enable debug capabilities for the entity. Often this is implemented by blowing a fuse on a chip as fuses cannot be restored once blown. This claim is identified by Claim Key X+8.

3.7. Location (loc) Claim

The "loc" (location) claim is a CBOR-formatted object that describes the location of the device entity from which the attestation originates. It is identified by Claim Key X+10. It is comprised of an array of additional subclaims that represent the actual location coordinates (latitude, longitude and altitude). The location coordinate claims are consistent with the WGS84 coordinate system [WGS84]. In addition, a subclaim providing the estimated accuracy of the location measurement is defined.

3.7.1. lat (latitude) claim

The "lat" (latitude) claim contains the value of the device location corresponding to its latitude coordinate. It is of data type FloatOrNumber and identified by Claim Key X+11.

3.7.2. long (longitude) claim

The "long" (longitude) claim contains the value of the device location corresponding to its longitude coordinate. It is of data type FloatOrNumber and identified by Claim Key X+12.

3.7.3. alt (altitude) claim

The "alt" (altitude) claim contains the value of the device location corresponding to its altitude coordinate (if available). It is of data type FloatOrNumber and identified by Claim Key X+13.

3.7.4. acc (accuracy) claim

The "acc" (accuracy) claim contains a value that describes the location accuracy. It is non-negative and expressed in meters. It is of data type FloatOrNumber and identified by Claim Key X+14.

3.7.5. altacc (altitiude accuracy) claim

The "altacc" (altitude accuracy) claim contains a value that describes the altitude accuracy. It is non-negative and expressed in meters. It is of data type FloatOrNumber and identified by Claim Key X+15.

3.7.6. heading claim

The "heading" claim contains a value that describes direction of motion for the entity. Its value is specified in degrees, between 0 and 360. It is of data type FloatOrNumber and identified by Claim Key X+16.

3.7.7. speed claim

The "speed" claim contains a value that describes the velocity of the entity in the horizontal direction. Its value is specified in meters/second and must be non-negative. It is of data type FloatOrNumber and identified by Claim Key X+17.

3.8. ts (timestamp) claim

The "ts" (timestamp) claim contains a timestamp derived using the same time reference as is used to generate an "iat" claim (see Section 3.1.6 of [RFC8392]). It is of the same type as "iat" (integer or floating-point), and is identified by Claim Key X+18. It is meant to designate the time at which a measurement was taken, when a location was obtained, or when a token was actually transmitted. The timestamp would be included as a subclaim under the "submod" or "loc" claims (in addition to the existing respective subclaims), or at the device level.

3.9. age claim

The "age" claim contains a value that represents the number of seconds that have elapsed since the token was created, measurement was made, or location was obtained. Typical attestable values are sent as soon as they are obtained. However in the case that such a value is buffered and sent at a later time and a sufficiently accurate time reference is unavailable for creation of a timestamp, then the age claim is provided. It is identified by Claim Key X+19.

3.10. uptime claim

The "uptime" claim contains a value that represents the number of seconds that have elapsed since the entity or submod was last booted. It is identified by Claim Key X+20.

3.11. The submods Claim

Some devices are complex, having many subsystems or submodules. A mobile phone is a good example. It may have several connectivity submodules for communications (e.g., WiFi and celluar). It may have sub systems for low-power audio and video playback. It may have one or more security-oriented subsystems like a TEE or a Secure Element.

The claims for each these can be grouped together in a submodule.

Specifically, the "submods" claim is an array. Each item in the array is a CBOR map containing all the claims for a particular submodule. It is identified by Claim Key X+22.

The security level of the submod is assumed to be at the same level as the main entity unless there is a security level claim in that submodule indicating otherwise. The security level of a submodule can never be higher (more secure) than the security level of the EAT it is a part of.

3.11.1. The submod_name Claim

Each submodule should have a submod_name claim that is descriptive name. This name should be the CBOR txt type.

3.11.2. Nested EATs, the eat Claim

It is allowed for one EAT to be embedded in another. This is for complex devices that have more than one subsystem capable of generating an EAT. Typically one will be the device-wide EAT that is low to medium security and another from a Secure Element or similar that is high security.

The contents of the "eat" claim must be a fully signed, optionally encrypted, EAT token. It is identified by Claim Key X+23.

4. IANA Considerations

4.1. Reuse of CBOR Web Token (CWT) Claims Registry

Claims defined for EAT are compatible with those of CWT so the CWT Claims Registry is re used. New new IANA registry is created. All EAT claims should be registered in the CWT Claims Registry.

4.1.1. Claims Registered by This Document

- Claim Name: UEID
- Claim Description: The Universal Entity ID
- JWT Claim Name: N/A
- Claim Key: X
- Claim Value Type(s): byte string
- Change Controller: IESG
- Specification Document(s): *this document*

TODO: add the rest of the claims in here

4.2. EAT CBOR Tag Registration

How an EAT consumer determines whether received CBOR-formatted data actually represents a valid EAT is application-dependent, much like a CWT. For instance, a specific MIME type associated with the EAT such as "application/eat" could be sufficient for identification of the EAT. Note however that EAT's can include other EAT's (e.g. a device EAT comprised of several submodule EAT's). In this case, a CBOR tag dedicated to the EAT will be required at least for the submodule EAT's and the tag must be a valid CBOR tag. In other words - the EAT CBOR tag can optionally prefix a device-level EAT, but a EAT CBOR tag must always prefix a submodule EAT. The proposed EAT CBOR tag is 71.

4.2.1. Tag Registered by This Document

- CBOR Tag: 71
- Data Item: Entity Attestation Token (EAT)
- Semantics: Entity Attestation Token (CWT), as defined in *this_doc*
- Reference: *this_doc*
- Point of Contact: Giridhar Mandyam, mandyam@qti.qualcomm.com

5. Privacy Considerations

Certain EAT claims can be used to track the owner of an entity and therefore implementations should consider providing privacy-preserving options dependent on the intended usage of the EAT. Examples would include suppresion of location claims for EAT's provided to unauthenticated consumers.

5.1. UEID Privacy Considerations

A UEID is usually not privacy preserving. Any set of relying parties that receives tokens that happen to be from a single device will be able to know the tokens are all from the same device and be able to track the device. Thus, in many usage situations ueid violates governmental privacy regulation. In other usage situations UEID will not be allowed for certain products like browsers that give privacy for the end user. it will often be the case that tokens will not have a UEID for these reasons.

There are several strategies that can be used to still be able to put UEID's in tokens:

- The device obtains explicit permission from the user of the device to use the UEID. This may be through a prompt. It may also be through a license agreement. For example, agreements for some online banking and brokerage services might already cover use of a UEID.

- The UEID is used only in a particular context or particular use case. It is used only by one relying party.

- The device authenticates the relying party and generates a derivate UEID just for that particular relying party. For example, the relying party could prove their identity cryptographically to the device, then the device generates a UEID just for that relying party by hashing a proofed relying party ID with the main device UEID.

Note that some of these privacy preservation strategies result in multiple UEIDs per device. Each UEID is used in a different context, use case or system on the device. However, from the view of the relying party, there is just one UEID and it is still globally universal across manufacturers.

6. Security Considerations

TODO: Perhpas this can be the same as CWT / COSE, but not sure yet because it involves so much entity / device security that those do not.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC7049] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, DOI 10.17487/RFC7049, October 2013, <https://www.rfc-editor.org/info/rfc7049>.

[RFC7519] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <https://www.rfc-editor.org/info/rfc7519>.

[RFC8152] Schaad, J., "CBOR Object Signing and Encryption (COSE)", RFC 8152, DOI 10.17487/RFC8152, July 2017, <https://www.rfc-editor.org/info/rfc8152>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8392] Jones, M., Wahlstroem, E., Erdtman, S., and H. Tschofenig, "CBOR Web Token (CWT)", RFC 8392, DOI 10.17487/RFC8392, May 2018, <https://www.rfc-editor.org/info/rfc8392>.

[TIME_T] The Open Group Base Specifications, "Vol. 1: Base Definitions, Issue 7", Section 4.15 'Seconds Since the Epoch', IEEE Std 1003.1, 2013 Edition, 2013, <http://pubs.opengroup.org/onlinepubs/9699919799/basedefs/V1_chap04.html#tag_04_15>.

[WGS84] National Imagery and Mapping Agency, "National Imagery and Mapping Agency Technical Report 8350.2, Third Edition", 2000, <http://earth-info.nga.mil/GandG/publications/tr8350.2/wgs84fin.pdf>.

7.2. Informative References

[ASN.1] International Telecommunication Union, "Information Technology -- ASN.1 encoding rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER)", ITU-T Recommendation X.690, 1994.

[IDevID] "IEEE Standard, "IEEE 802.1AR Secure Device Identifier"", December 2009, <http://standards.ieee.org/findstds/standard/802.1AR-2009.html>.

[Webauthn] Worldwide Web Consortium, "Web Authentication: A Web API for accessing scoped credentials", 2016.

Appendix A. Examples

A.1. Very Simple EAT

This is shown in CBOR diagnostic form. Only the payload signed by COSE is shown.

```
{
   / nonce /                  11:h'948f8860d13a463e8e',
   / UEID /                    8:h'0198f50a4ff6c05861c8860d13a638ea4fe2f',
   / secbootenabled /         13:true,
   / debugpermanentdisable / 15:true,
   / ts /                     21:1526542894,
}
```

A.2. Example with Submodules, Nesting and Security Levels

```
{
   / nonce /                  11:h'948f8860d13a463e8e',
   / UEID /                    8:h'0198f50a4ff6c05861c8860d13a638ea4fe2f',
   / secbootenabled /         13:true,
   / debugpermanentdisable / 15:true,
   / ts /                     21:1526542894,
   / seclevel /               10:3, / secure restriced OS /

   / submods / 30:
      [
         / 1st submod, an Android Application / {
           / submod_name /   30:'Android App "Foo"',
           / seclevel /      10:1, / unrestricted /
           / app data /  -70000:'text string'
         },
         / 2nd submod, A nested EAT from a secure element / {
           / submod_name / 30:'Secure Element EAT',
           / eat /         31:71( 18(
              / an embedded EAT / [ /...COSE_Sign1 bytes with payload.../ ]
                           ))
         }
         / 3rd submod, information about Linux Android / {
            / submod_name/ 30:'Linux Android',
            / seclevel /   10:1, / unrestricted /
            / custom - release / -80000:'8.0.0',
            / custom - version / -80001:'4.9.51+'
         }
      ]
}
```

Authors' Addresses

Giridhar Mandyam
Qualcomm Technologies Inc.
5775 Morehouse Drive
San Diego, California
USA

Phone: +1 858 651 7200
EMail: mandyam@qti.qualcomm.com

Laurence Lundblade
Qualcomm Technologies Inc.
5775 Morehouse Drive
San Diego, California
USA

Phone: +1 858 658 3584
EMail: llundbla@qti.qualcomm.com

Miguel Ballesteros
Qualcomm Technologies Inc.
5775 Morehouse Drive
San Diego, California
USA

Phone: +1 858 651 4299
EMail: mballest@qti.qualcomm.com

Jeremy O'Donoghue
Qualcomm Technologies Inc.
279 Farnborough Road
Farnborough GU14 7LS
United Kingdom

Phone: +44 1252 363189
EMail: jodonogh@qti.qualcomm.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 23, 2019

G. Mandyam
Qualcomm Technologies Inc.
L. Lundblade
Security Theory LLC
M. Ballesteros
J. O'Donoghue
Qualcomm Technologies Inc.

November 19, 2018

The Entity Attestation Token (EAT)
draft-mandyam-eat-01

Abstract

An attestation format based on concise binary object representation (CBOR) is proposed that is suitable for inclusion in a CBOR Web Token (CWT), know as the Entity Attestation Token (EAT). The associated data can be used by a relying party to assess the security state of a remote device or module.

Contributing

TBD

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Remote device attestation is fundamental service that allows a remote device such as a mobile phone, an Internet-of-Things (IoT) device, or other endpoint to prove itself to a relying party, a server or a service. This allows the relying party to know some characteristics about the device and decide whether it trusts the device.

Remote attestation is a fundamental service that can underlie other protocols and services that need to know about the trustworthiness of the device before proceeding. One good example is biometric authentication where the biometric matching is done on the device. The relying party needs to know that the device is one that is known to do biometric matching correctly. Another example is content protection where the relying party wants to know the device will protect the data. This generalizes on to corporate enterprises that might want to know that a device is trustworthy before allowing corporate data to be accessed by it.

The notion of attestation here is large and may include, but is not limited to the following:

- Proof of the make and model of the device hardware (HW)

- o Proof of the make and model of the device processor, particularly for security oriented chips

- o Measurement of the software (SW) running on the device

- o Configuration and state of the device

- o Environmental characteristics of the device such as its GPS location

The required data format should be general purpose and extensible so that it can work across many use cases. This is why CBOR (see [RFC7049]) was chosen as the format -- it already supports a rich set of data types, and is both expressive and extensible. It translates well to JSON for good interoperation with web technology. It is compact and can work on very small IoT device. The format proposed here is small enough that a limited version can be implemented in pure hardware gates with no software at all. Moreover, the attestation data is defined in the form of claims that is the same as CBOR Web Token (CWT, see [RFC8392]). This is the motivation for defining the Entity Attestation Token, i.e. EAT.

1.1. Entity Overview

An "entity" can be any device or device subassembly ("submodule") that can generate its own attestation in the form of an EAT. The attestation should be cryptographically verifiable by the EAT consumer. An EAT at the device-level can be composed of several submodule EAT's. It is assumed that any entity that can create an EAT does so by means of a dedicated root-of-trust (RoT).

Modern devices such as a mobile phone have many different execution environments operating with different security levels. For example it is common for a mobile phone to have an "apps" environment that runs an operating system (OS) that hosts a plethora of downloadable apps. It may also have a TEE (Trusted Execution Environment) that is distinct, isolated, and hosts security-oriented functionality like biometric authentication. Additionally it may have an eSE (embedded Secure Element) - a high security chip with defenses against HW attacks that can serve as a RoT. This device attestation format allows the attested data to be tagged at a security level from which it originates. In general, any discrete execution environment that has an identifiable security level can be considered an entity.

1.2. Use of CBOR and COSE

Fundamentally this attestation format is a verifiable data format. It is a collection of data items that can be signed by an attestation key, hashed, and/or encrypted. As per Section 7 of [RFC8392], the verification method is in the CWT using the CBOR Object Signing and Encryption (COSE) methodology (see [RFC8152]).

In addition, the reported attestation data could be determined within the secure operating environment or written to it from an external and presumably less trusted entity on the device. In either case, the source of the reported data must be identifiable by the relying party.

This attestation format is a single relatively simple signed message. It is designed to be incorporated into many other protocols and many other transports. It is also designed such that other SW and apps can add their own data to the message such that it is also attested.

1.3. EAT Operating Models

At least the following three participants exist in all EAT operating models. Some operating models have additional participants.

The Entity. This is the phone, the IoT device, the sensor, the sub-assembly or such that the attestation provides information about.

The Manufacturer. The company that made the entity. This may be a chip vendor, a circuit board module vendor or a vendor of finished consumer products.

The Relying Party. The server, service or company that makes use of the information in the EAT about the entity.

In all operating models, the manufacturer provisions some secret attestation key material (AKM) into the entity during manufacturing. This might be during the manufacturer of a chip at a fabrication facility (fab) or during final assembly of a consumer product or any time in between. This attestation key material is used for signing EATs.

In all operating models, hardware and/or software on the entity create an EAT of the format described in this document. The EAT is always signed by the attestation key material provisioned by the manufacturer.

In all operating models, the relying party must end up knowing that the signature on the EAT is valid and consistent with data from claims in the EAT. This can happen in many different ways. Here are some examples.

- The EAT is transmitted to the relying party. The relying party gets corresponding key material (e.g. a root certificate) from the manufacturer. The relying party performs the verification.

- The EAT is transmitted to the relying party. The relying party transmits the EAT to a verification service offered by the manufacturer. The server returns the validated claims.

- The EAT is transmitted directly to a verification service, perhaps operated by the manufacturer or perhaps by another party. It verifies the EAT and makes the validated claims available to the relying party. It may even modify the claims in some way and re-sign the EAT (with a different signing key).

This standard supports all these operating models and does not prefer one over the other. It is important to support this variety of operating models to generally facilitate deployment and to allow for some special scenarios. One special scenario has a validation service that is monetized, most likely by the manufacturer. In another, a privacy proxy service processes the EAT before it is transmitted to the relying party. In yet another, symmetric key material is used for signing. In this case the manufacturer should perform the verification, because any release of the key material would enable a participant other than the entity to create valid signed EATs.

1.4. What is Not Standardized

1.4.1. Transmission Protocol

EATs may be transmitted by any protocol. For example, they might be added in extension fields of other protocols, bundled into an HTTP header, or just transmitted as files. This flexibility is intentional to allow broader adoption. This flexibility is possible because EAT's are self-secured with signing (and possibly additionally with encryption and anti-replay). The transmission protocol is not required to fulfill any additional security requirements.

For certain devices, a direct connection may not exist between the EAT-producing device and the Relying Party. In such cases, the EAT should be protected against malicious access. The use of COSE allows for signing and encryption of the EAT. Therefore even if the EAT is conveyed through intermediaries between the device and Relying Party, such intermediaries cannot easily modify the EAT payload or alter the signature.

1.4.2. Signing Scheme

The term "signing scheme" is used to refer to the system that includes end-end process of establishing signing attestation key material in the entity, signing the EAT, and verifying it. This might involve key IDs and X.509 certificate chains or something similar but different. The term "signing algorithm" refers just to the algorithm ID in the COSE signing structure. No particular signing algorithm or signing scheme is required by this standard.

There are three main implementation issues driving this. First, secure non-volatile storage space in the entity for the attestation key material may be highly limited, perhaps to only a few hundred bits, on some small IoT chips. Second, the factory cost of provisioning key material in each chip or device may be high, with even millisecond delays adding to the cost of a chip. Third, privacy-preserving signing schemes like ECDAA (Elliptic Curve Direct Anonymous Attestation) are complex and not suitable for all use cases.

Eventually some form of standardization of the signing scheme may be required. This might come in the form of another standard that adds to this document, or when there is clear convergence on a small number of signing schemes this standard can be updated.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

This document reuses terminology from JWT [RFC7519], COSE [RFC8152], and CWT [RFC8392].

StringOrURI. The "StringOrURI" term in this specification has the same meaning and processing rules as the JWT "StringOrURI" term defined in Section 2 of [RFC7519], except that it is represented as a CBOR text string instead of a JSON text string.

NumericDate. The "NumericDate" term in this specification has the same meaning and processing rules as the JWT "NumericDate" term defined in Section 2 of [RFC7519], except that it is represented as a CBOR numeric date (from Section 2.4.1 of [RFC7049]) instead of a JSON number. The encoding is modified so that the leading tag 1 (epoch-based date/time) MUST be omitted.

Claim Name. The human-readable name used to identify a claim.

Claim Key. The CBOR map key used to identify a claim.

Claim Value. The CBOR map value representing the value of the claim.

CWT Claims Set. The CBOR map that contains the claims conveyed by the CWT.

FloatOrNumber. The "FloatOrNumber" term in this specification is the type of a claim that is either a CBOR positive integer, negative integer or floating point number.

Attestation Key Material (AKM). The key material used to sign the EAT token. If it is done symmetrically with HMAC, then this is a simple symmetric key. If it is done with ECC, such as an IEEE DevID [IDevID], then this is the private part of the EC key pair. If ECDAA is used, (e.g., as used by Enhanced Privacy ID, i.e. EPID) then it is the key material needed for ECDAA.

3. The Claims

3.1. Universal Entity ID (UEID) Claim

UEID’s identify individual manufactured entities / devices such as a mobile phone, a water meter, a Bluetooth speaker or a networked security camera. It may identify the entire device or a submodule or subsystem. It does not identify types, models or classes of devices. It is akin to a serial number, though it does not have to be sequential.

It is identified by Claim Key X (X is TBD).

UEID’s must be universally and globally unique across manufacturers and countries. UEIDs must also be unique across protocols and systems, as tokens are intended to be embedded in many different protocols and systems. No two products anywhere, even in completely different industries made by two different manufacturers in two different countries. should have the same UEID (if they are not global and universal in this way then relying parties receiving them will have to track other characteristics of the device to keep devices distinct between manufacturers).

The UEID should be permanent. It should never change for a given device / entity. In addition, it should not be reprogrammable.

UEID's are binary byte-strings (resulting in a smaller size than text strings). When handled in text-based protocols, they should be base-64 encoded.

UEID's are variable length with a maximum size of 33 bytes (1 type byte and 256 bits). A receivers of a token with UEIDs may reject the token if a UEID is larger than 33 bytes.

UEID's are not designed for direct use by humans (e.g., printing on the case of a device), so no textual representation is defined.

A UEID is a byte string. From the consumer's view (the rely party) it is opaque with no bytes having any special meaning.

When the entity constructs the UEID, the first byte is a type and the following bytes the ID for that type. Several types are allowed to accommodate different industries and different manufacturing processes and to give options to avoid paying fees for certain types of manufacturer registrations.

Type Byte	Type Name	Specification
0x01	GUID	This is a 128 to 256 bit random number generated once and stored in the device. The GUID may be constructed from various identifiers on the device using a hash function or it may be just the raw random number. In any case, the random number must have entropy of at least 128 bits as this is what gives the global
0x02	IEEE EUI	This makes use of the IEEE company identification registry. An EUI is made up of an OUI and OUI-36 or a CID, different registered company identifiers, and some unique per-device identifier. EUIs are often the same as or similar to MAC addresses. (Note that while devices with multiple network interfaces may have multiple MAC addresses, there is only one UEID for a device) TODO: normative references to IEEE.
0x03	IMEI	TODO: figure how to specify IMEIs

Table 1: UEID Composition Types

The consumer (the Relying Party) of a UEID should treat a UEID as a completely opaque string of bytes and not make any use of its internal structure. For example they should not use the OUI part of a type 0x02 UEID to identify the manufacturer of the device. Instead they should use the OUI claim that is defined elsewhere. The reasons for this are:

- UEIDs types may vary freely from one manufacturer to the next.
- New types of UEIDs may be created. For example a type 0x04 UEID may be created based on some other manufacturer registration scheme.
- Device manufacturers are allowed to change from one type of UEID to another anytime they want. For example they may find they can optimize their manufacturing by switching from type 0x01 to type 0x02 or vice versa. The main requirement on the manufacturer is that UEIDs be universally unique.

3.2. Origination (origination) Claims

This claim describes the parts of the device or entity that are creating the EAT. Often it will be tied back to the device or chip manufacturer. The following table gives some examples:

Name	Description
Acme-TEE	The EATs are generated in the TEE authored and configured by "Acme"
Acme-TPM	The EATs are generated in a TPM manufactured by "Acme"
Acme-Linux-Kernel	The EATs are generated in a Linux kernel configured and shipped by "Acme"
Acme-TA	The EATs are generated in a Trusted Application (TA) authored by "Acme"

The claim is represented by Claim Key X+1. It is type StringOrURI.

TODO: consider a more structure approach where the name and the URI and other are in separate fields.

TODO: This needs refinement. It is somewhat parallel to issuer claim in CWT in that it describes the authority that created the token.

3.3. OEM identification by IEEE OUI

This claim identifies a device OEM by the IEEE OUI. Reference TBD. It is a byte string representing the OUI in binary form in network byte order (TODO: confirm details).

Companies that have more than one IEEE OUI registered with IEEE should pick one and prefer that for all their devices.

Note that the OUI is in common use as a part of MAC Address. This claim is only the first bits of the MAC address that identify the manufacturer. The IEEE maintains a registry for these in which many companies participate. This claim is represented by Claim Key TBD.

3.4. Security Level (seclevel) Claim

EATs have a claim that roughly characterizes the device / entities ability to defend against attacks aimed at capturing the signing key, forging claims and at forging EATs. This is done by roughly defining four security levels as described below. This is similar to the security levels defined in the Metadata Service definied by the Fast Identity Online (FIDO) Alliance (TODO: reference).

These claims describe security environment and countermeasures available on the end-entity / client device where the attestation key reside and the claims originate.

This claim is identified by Claim Key X+2. The value is an integer between 1 and 4 as defined below.

1 - Unrestricted There is some expectation that implementor will protect the attestation signing keys at this level. Otherwise the EAT provides no meaningful security assurances.

2- Restricted Entities at this level should not be general-purpose operating environments that host features such as app download systems, web browsers and complex productivity applications. It is akin to the Secure Restricted level (see below) without the security orientation. Examples include a WiFi subsystem, an IoT camera, or sensor device.

3 - Secure Restricted Entities at this level must meet the critera defined by FIDO Allowed Restricted Operating Environments (TODO: reference). Examples include TEE's and schemes using virtualization-based security. Like the FIDO security goal, security at this level is aimed at defending well against large-scale network / remote attacks against the device.

4 - Hardware Entities at this level must include substantial defense against physical or electrical attacks against the device itself. It is assumed any potential attacker has captured the device and can disassemble it. Example include TPMs and Secure Elements.

This claim is not intended as a replacement for a proper end-device security certification schemes such as those based on FIPS (TODO: reference) or those based on Common Criteria (TODO: reference). The claim made here is solely a self-claim made by the Entity Originator.

3.5. Nonce (nonce) Claim

The "nonce" (Nonce) claim represents a random value that can be used to avoid replay attacks. This would be ideally generated by the CWT consumer. This value is intended to be a CWT companion claim to the existing JWT claim **_IANAJWT_ (TODO: fix this reference). The nonce claim is identified by Claim Key X+3.

3.6. Secure Boot and Debug Enable State Claims

3.6.1. Secure Boot Enabled (secbootenabled) Claim

The "secbootenabled" (Secure Boot Enabled) claim represents a boolean value that indicates whether secure boot is enabled either for an entire device or an individual submodule. If it appears at the device level, then this means that secure boot is enabled for all submodules. Secure boot enablement allows a secure boot loader to authenticate software running either in a device or a submodule prior allowing execution. This claim is identified by Claim Key X+4.

3.6.2. Debug Disabled (debugdisabled) Claim

The "debugdisabled" (Debug Disabled) claim represents a boolean value that indicates whether debug capabilities are disabled for an entity (i.e. value of 'true'). Debug disablement is considered a prerequisite before an entity is considered operational. This claim is identified by Claim Key X+5.

3.6.3. Debug Disabled Since Boot (debugdisabledsincebboot) Claim

The "debugdisabledsinceboot" (Debug Disabled Since Boot) claim represents a boolean value that indicates whether debug capabilities for the entity were not disabled in any way since boot (i.e. value of 'true'). This claim is identified by Claim Key X+6.

3.6.4. Debug Permanent Disable (debugpermanentdisable) Claim

The "debugpermanentdisable" (Debug Permanent Disable) claim represents a boolean value that indicates whether debug capabilities for the entity are permanently disabled (i.e. value of 'true'). This value can be set to 'true' also if only the manufacturer is allowed to enabled debug, but the end user is not. This claim is identified by Claim Key X+7.

3.6.5. Debug Full Permanent Disable (debugfullpermanentdisable) Claim

The "debugfullpermanentdisable" (Debug Full Permanent Disable) claim represents a boolean value that indicates whether debug capabilities for the entity are permanently disabled (i.e. value of 'true'). This value can only be set to 'true' if no party can enable debug capabilities for the entity. Often this is implemented by blowing a fuse on a chip as fuses cannot be restored once blown. This claim is identified by Claim Key X+8.

3.7. Location (loc) Claim

The "loc" (location) claim is a CBOR-formatted object that describes the location of the device entity from which the attestation originates. It is identified by Claim Key X+10. It is comprised of an array of additional subclaims that represent the actual location coordinates (latitude, longitude and altitude). The location coordinate claims are consistent with the WGS84 coordinate system [WGS84]. In addition, a subclaim providing the estimated accuracy of the location measurement is defined.

3.7.1. lat (latitude) claim

The "lat" (latitude) claim contains the value of the device location corresponding to its latitude coordinate. It is of data type FloatOrNumber and identified by Claim Key X+11.

3.7.2. long (longitude) claim

The "long" (longitude) claim contains the value of the device location corresponding to its longitude coordinate. It is of data type FloatOrNumber and identified by Claim Key X+12.

3.7.3. alt (altitude) claim

The "alt" (altitude) claim contains the value of the device location corresponding to its altitude coordinate (if available). It is of data type FloatOrNumber and identified by Claim Key X+13.

3.7.4. acc (accuracy) claim

The "acc" (accuracy) claim contains a value that describes the location accuracy. It is non-negative and expressed in meters. It is of data type FloatOrNumber and identified by Claim Key X+14.

3.7.5. altacc (altitude accuracy) claim

The "altacc" (altitude accuracy) claim contains a value that describes the altitude accuracy. It is non-negative and expressed in meters. It is of data type FloatOrNumber and identified by Claim Key X+15.

3.7.6. heading claim

The "heading" claim contains a value that describes direction of motion for the entity. Its value is specified in degrees, between 0 and 360. It is of data type FloatOrNumber and identified by Claim Key X+16.

3.7.7. speed claim

The "speed" claim contains a value that describes the velocity of the entity in the horizontal direction. Its value is specified in meters/second and must be non-negative. It is of data type FloatOrNumber and identified by Claim Key X+17.

3.8. ts (timestamp) claim

The "ts" (timestamp) claim contains a timestamp derived using the same time reference as is used to generate an "iat" claim (see Section 3.1.6 of [RFC8392]). It is of the same type as "iat" (integer or floating-point), and is identified by Claim Key X+18. It is meant to designate the time at which a measurement was taken, when a location was obtained, or when a token was actually transmitted. The timestamp would be included as a subclaim under the "submod" or "loc" claims (in addition to the existing respective subclaims), or at the device level.

3.9. age claim

The "age" claim contains a value that represents the number of seconds that have elapsed since the token was created, measurement was made, or location was obtained. Typical attestable values are sent as soon as they are obtained. However in the case that such a value is buffered and sent at a later time and a sufficiently accurate time reference is unavailable for creation of a timestamp, then the age claim is provided. It is identified by Claim Key X+19.

3.10. uptime claim

The "uptime" claim contains a value that represents the number of seconds that have elapsed since the entity or submod was last booted. It is identified by Claim Key X+20.

3.11. The submods Claim

Some devices are complex, having many subsystems or submodules. A mobile phone is a good example. It may have several connectivity submodules for communications (e.g., WiFi and cellular). It may have sub systems for low-power audio and video playback. It may have one or more security-oriented subsystems like a TEE or a Secure Element.

The claims for each these can be grouped together in a submodule.

Specifically, the "submods" claim is an array. Each item in the array is a CBOR map containing all the claims for a particular submodule. It is identified by Claim Key X+22.

The security level of the submod is assumed to be at the same level as the main entity unless there is a security level claim in that submodule indicating otherwise. The security level of a submodule can never be higher (more secure) than the security level of the EAT it is a part of.

3.11.1. The submod_name Claim

Each submodule should have a submod_name claim that is descriptive name. This name should be the CBOR txt type.

3.11.2. Nested EATs, the eat Claim

It is allowed for one EAT to be embedded in another. This is for complex devices that have more than one subsystem capable of generating an EAT. Typically one will be the device-wide EAT that is low to medium security and another from a Secure Element or similar that is high security.

The contents of the "eat" claim must be a fully signed, optionally encrypted, EAT token. It is identified by Claim Key X+23.

4. CBOR Interoperability

EAT is a one-way protocol. It only defines a single message that goes from the entity to the server. The entity implementation will often be in a contained environment with little RAM and the server will usually not be. The following requirements for interoperability take that into account. The entity can generally use whatever encoding it wants. The server is required to support just about every encoding.

Canonical CBOR encoding is explicitly NOT required as it would place an unnecessary burden on the entity implementation.

4.1. Integer Encoding (major type 0 and 1)

The entity may use any integer encoding allowed by CBOR. The server MUST accept all integer encodings allowed by CBOR.

4.2. String Encoding (major type 2 and 3)

The entity can use any string encoding allowed by CBOR including indefinite lengths. It may also encode the lengths of strings in any way allowed by CBOR. The server must accept all string encodings.

Major type 2, bstr, SHOULD be have tag 21, 22 or 23 to indicate conversion to base64 or such when converting to JSON.

4.3. Map and Array Encoding (major type 4 and 5)

The entity can use any array or map encoding allowed by CBOR including indefinite lengths. Sorting of map keys is not required. Duplicate map keys are not allowed. The server must accept all array and map encodings. The server may reject maps with duplicate map keys.

4.4. Date and Time

The entity should send dates as tag 1 encoded as 64-bit or 32-bit integers. The entity may not send floating point dates. The server must support tag 1 epoch based dates encoded as 64-bit or 32-bit integers.

The entity may send tag 0 dates, however tag 1 is preferred. The server must support tag 0 UTC dates.

4.5. URIs

URIs should be encoded as text strings and marked with tag 32.

4.6. Floating Point

Encoding data in floating point is to be used only if necessary. Location coordinates are always in floating point. The server must support decoding of all types of floating point.

4.7. Other types

Use of Other types like bignums, regular expressions and so SHOULD NOT be used. The server MAY support them, but is not required to. Use of these tags is

5. IANA Considerations

5.1. Reuse of CBOR Web Token (CWT) Claims Registry

Claims defined for EAT are compatible with those of CWT so the CWT Claims Registry is re used. New new IANA registry is created. All EAT claims should be registered in the CWT Claims Registry.

5.1.1. Claims Registered by This Document

- Claim Name: UEID
- Claim Description: The Universal Entity ID
- JWT Claim Name: N/A
- Claim Key: X
- Claim Value Type(s): byte string
- Change Controller: IESG
- Specification Document(s): *this document*

TODO: add the rest of the claims in here

5.2. EAT CBOR Tag Registration

How an EAT consumer determines whether received CBOR-formatted data actually represents a valid EAT is application-dependent, much like a CWT. For instance, a specific MIME type associated with the EAT such as "application/eat" could be sufficient for identification of the EAT. Note however that EAT's can include other EAT's (e.g. a device EAT comprised of several submodule EAT's). In this case, a CBOR tag dedicated to the EAT will be required at least for the submodule EAT's and the tag must be a valid CBOR tag. In other words - the EAT CBOR tag can optionally prefix a device-level EAT, but a EAT CBOR tag must always prefix a submodule EAT. The proposed EAT CBOR tag is 71.

5.2.1. Tag Registered by This Document

- CBOR Tag: 71
- Data Item: Entity Attestation Token (EAT)
- Semantics: Entity Attestation Token (CWT), as defined in *this_doc*

- o Reference: *this_doc*

- o Point of Contact: Giridhar Mandyam, mandyam@qti.qualcomm.com

6. Privacy Considerations

Certain EAT claims can be used to track the owner of an entity and therefore implementations should consider providing privacy-preserving options dependent on the intended usage of the EAT. Examples would include suppression of location claims for EAT's provided to unauthenticated consumers.

6.1. UEID Privacy Considerations

A UEID is usually not privacy preserving. Any set of relying parties that receives tokens that happen to be from a single device will be able to know the tokens are all from the same device and be able to track the device. Thus, in many usage situations ueid violates governmental privacy regulation. In other usage situations UEID will not be allowed for certain products like browsers that give privacy for the end user. it will often be the case that tokens will not have a UEID for these reasons.

There are several strategies that can be used to still be able to put UEID's in tokens:

- o The device obtains explicit permission from the user of the device to use the UEID. This may be through a prompt. It may also be through a license agreement. For example, agreements for some online banking and brokerage services might already cover use of a UEID.

- o The UEID is used only in a particular context or particular use case. It is used only by one relying party.

- o The device authenticates the relying party and generates a derived UEID just for that particular relying party. For example, the relying party could prove their identity cryptographically to the device, then the device generates a UEID just for that relying party by hashing a proofed relying party ID with the main device UEID.

Note that some of these privacy preservation strategies result in multiple UEIDs per device. Each UEID is used in a different context, use case or system on the device. However, from the view of the relying party, there is just one UEID and it is still globally universal across manufacturers.

7. Security Considerations

TODO: Perhaps this can be the same as CWT / COSE, but not sure yet because it involves so much entity / device security that those do not.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC7049] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, DOI 10.17487/RFC7049, October 2013, <https://www.rfc-editor.org/info/rfc7049>.

[RFC7519] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <https://www.rfc-editor.org/info/rfc7519>.

[RFC8152] Schaad, J., "CBOR Object Signing and Encryption (COSE)", RFC 8152, DOI 10.17487/RFC8152, July 2017, <https://www.rfc-editor.org/info/rfc8152>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8392] Jones, M., Wahlstroem, E., Erdtman, S., and H. Tschofenig, "CBOR Web Token (CWT)", RFC 8392, DOI 10.17487/RFC8392, May 2018, <https://www.rfc-editor.org/info/rfc8392>.

[TIME_T] The Open Group Base Specifications, "Vol. 1: Base Definitions, Issue 7", Section 4.15 'Seconds Since the Epoch', IEEE Std 1003.1, 2013 Edition, 2013, <http://pubs.opengroup.org/onlinepubs/9699919799/basedefs/V1_chap04.html#tag_04_15>.

[WGS84] National Imagery and Mapping Agency, "National Imagery and Mapping Agency Technical Report 8350.2, Third Edition", 2000, <http://earth-info.nga.mil/GandG/publications/tr8350.2/wgs84fin.pdf>.

8.2. Informative References

[ASN.1] International Telecommunication Union, "Information Technology -- ASN.1 encoding rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER)", ITU-T Recommendation X.690, 1994.

[IDevID] "IEEE Standard, "IEEE 802.1AR Secure Device Identifier"", December 2009, <http://standards.ieee.org/findstds/standard/802.1AR-2009.html>.

[Webauthn] Worldwide Web Consortium, "Web Authentication: A Web API for accessing scoped credentials", 2016.

Appendix A. Examples

A.1. Very Simple EAT

This is shown in CBOR diagnostic form. Only the payload signed by COSE is shown.

```
{
   / nonce /                  11:h'948f8860d13a463e8e',
   / UEID /                    8:h'0198f50a4ff6c05861c8860d13a638ea4fe2f',
   / secbootenabled /         13:true,
   / debugpermanentdisable / 15:true,
   / ts /                     21:1526542894,
}
```

A.2. Example with Submodules, Nesting and Security Levels

```
{
   / nonce /                  11:h'948f8860d13a463e8e',
   / UEID /                    8:h'0198f50a4ff6c05861c8860d13a638ea4fe2f',
   / secbootenabled /         13:true,
   / debugpermanentdisable / 15:true,
   / ts /                     21:1526542894,
   / seclevel /               10:3, / secure restriced OS /

   / submods / 30:
      [
         / 1st submod, an Android Application / {
           / submod_name /   30:'Android App "Foo"',
           / seclevel /      10:1, / unrestricted /
           / app data /  -70000:'text string'
         },
         / 2nd submod, A nested EAT from a secure element / {
           / submod_name / 30:'Secure Element EAT',
           / eat /         31:71( 18(
              / an embedded EAT / [ /...COSE_Sign1 bytes with payload.../ ]
                          ))
         }
         / 3rd submod, information about Linux Android / {
            / submod_name/ 30:'Linux Android',
            / seclevel /   10:1, / unrestricted /
            / custom - release / -80000:'8.0.0',
            / custom - version / -80001:'4.9.51+'
         }
      ]
}
```

Authors' Addresses

Giridhar Mandyam
Qualcomm Technologies Inc.
5775 Morehouse Drive
San Diego, California
USA

Phone: +1 858 651 7200
EMail: mandyam@qti.qualcomm.com

Laurence Lundblade
Security Theory LLC

EMail: lgl@island-resort.com

Miguel Ballesteros
Qualcomm Technologies Inc.
5775 Morehouse Drive
San Diego, California
USA

Phone: +1 858 651 4299
EMail: mballest@qti.qualcomm.com

Jeremy O'Donoghue
Qualcomm Technologies Inc.
279 Farnborough Road
Farnborough GU14 7LS
United Kingdom

Phone: +44 1252 363189
EMail: jodonogh@qti.qualcomm.com