SAND2021-6730C

Sandia National Laboratories

Exceptional service in the national interest

FARM

Malware Information Sharing

Ken Chiang

June 8th, 2021

Automated Malware Analysis Symposium

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

U.S. DEPARTMENT OF ENERGY

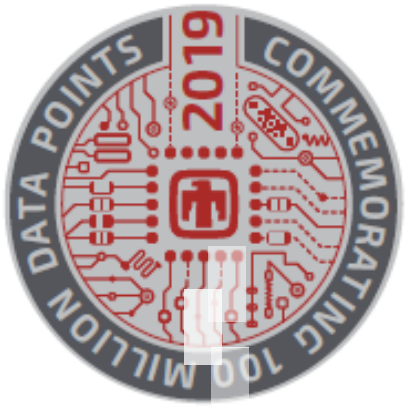
NNSA

Agenda

- FARM background - 5 min
- Federation concept – 10 min
- Under the hood: Latest version (FARM6) architecture - 5min
- Q & A – 10min

FARM Background

- Forensic Analysis Repository for Malware
  - First version circa ~2006
  - Sandia LDRD funded
- 121 million samples in current version FARM5
- Anticipated growth rate 100k-300k unique samples per day in FARM6.
  - Sandia IR
  - public/private malware feeds
  - DOE network files
- Enrichment sources:
  - Public/Private IOC lists
  - Public/Private Yara sigs
  - Virus Total
  - US Cert Reports
- Goal:
  - Share malware information across .gov to:
    - Minimize **triage** time
    - Increase probability of **attribution**



FARM Background – Capabilities

- Capabilities (COTS/GOTS/Sandia Developed)

STATIC ANALYSIS

- **Malware ID**
- **Related Graph**
- Signify
- Exiftool
- FireEye Floss, Capa
- Antivirus
- Yara

UNPACKERS

- Titanium Core 2
- Memory Carving
- UPX

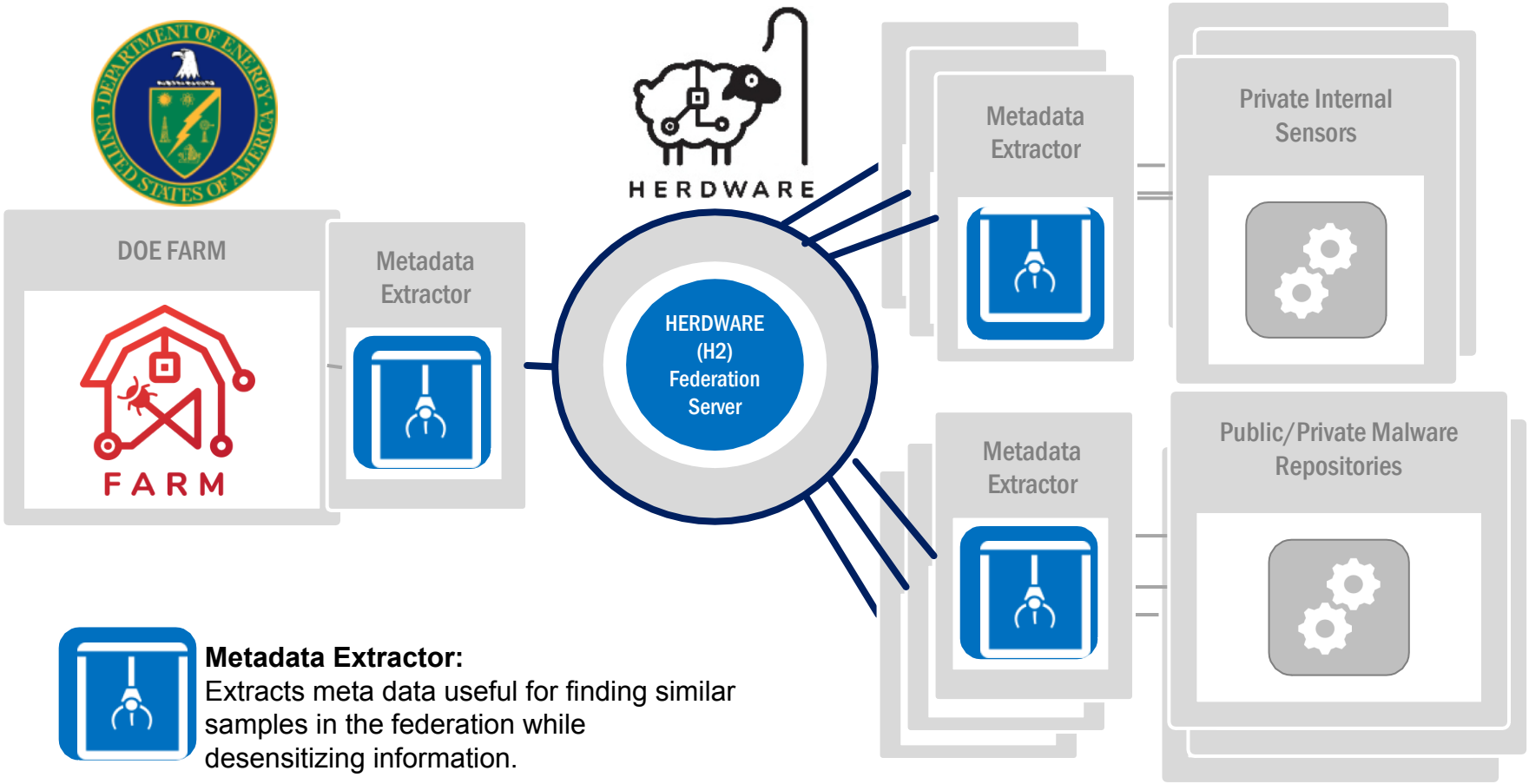
DYNAMIC ANALYSIS

- Disk/Network/Process/Memory/System calls
- Virtual KVM
- Bare metal

MACHINE LEARNING

- Syscall avatar
- Syscall ensemble
- PE avatar
- **BIFROST**

Federation Concept - Architecture



Metadata Extractor:
Extracts meta data useful for finding similar samples in the federation while desensitizing information.

Federation Goal:
Allow for sharing of attribution information across repository boundaries

FARM participates in the Federation with features and labels

Label Sources:

- IR cases
- US-CERT reports
- RE efforts

Types:

- Malware family
- Campaign
- Actor

Propagated using relationships:

- Code similarity
- Unpacking
- Static/Dynamic Traits

Confidence Scoring based on:

- Features
- Static traits
- Dynamic behaviors
- Inferred Traits

Federation Concept

Current Status:

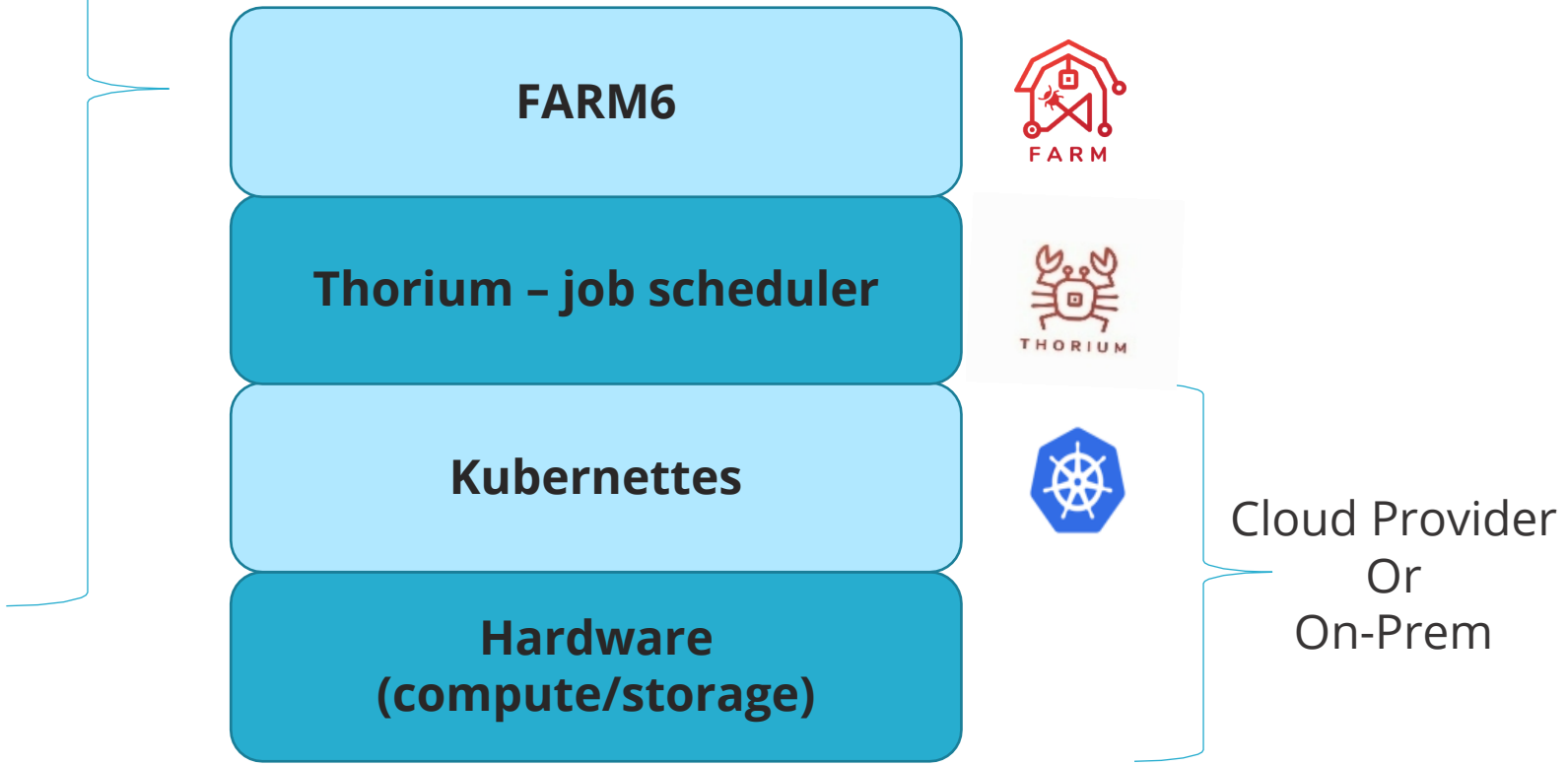
- Pilot server at SNL
- Client docker image available
- File type supported currently:
  - EXE (x86_64)
  - PDF(javascript)
  - OFFICE (Vbscript)
- FARM contributes
  - TLP white/green and public features and propagated labels

Others can participate by:

- Anonymous query using desensitized extracted features
- Contribute features
- Contribute labels
  - Actors
  - Malware Family
  - Campaign
  - Other?

FARM6 architecture

- API/GUI
- S3 interface
- Dynamic analysis
- Elasticsearch
- Plugin architecture
- Triage workflows



FARM – Further information sharing

- FARM – https://farm.sandia.gov
  - Contact farm@sandia.gov for access
- FARM6 (Open Beta)
  - Government use license on source code
- Federation
  - Anonymous query
  - Manual contribution of malware features and labels
  - Automate contribution of malware features and labels

FARM – Q&A