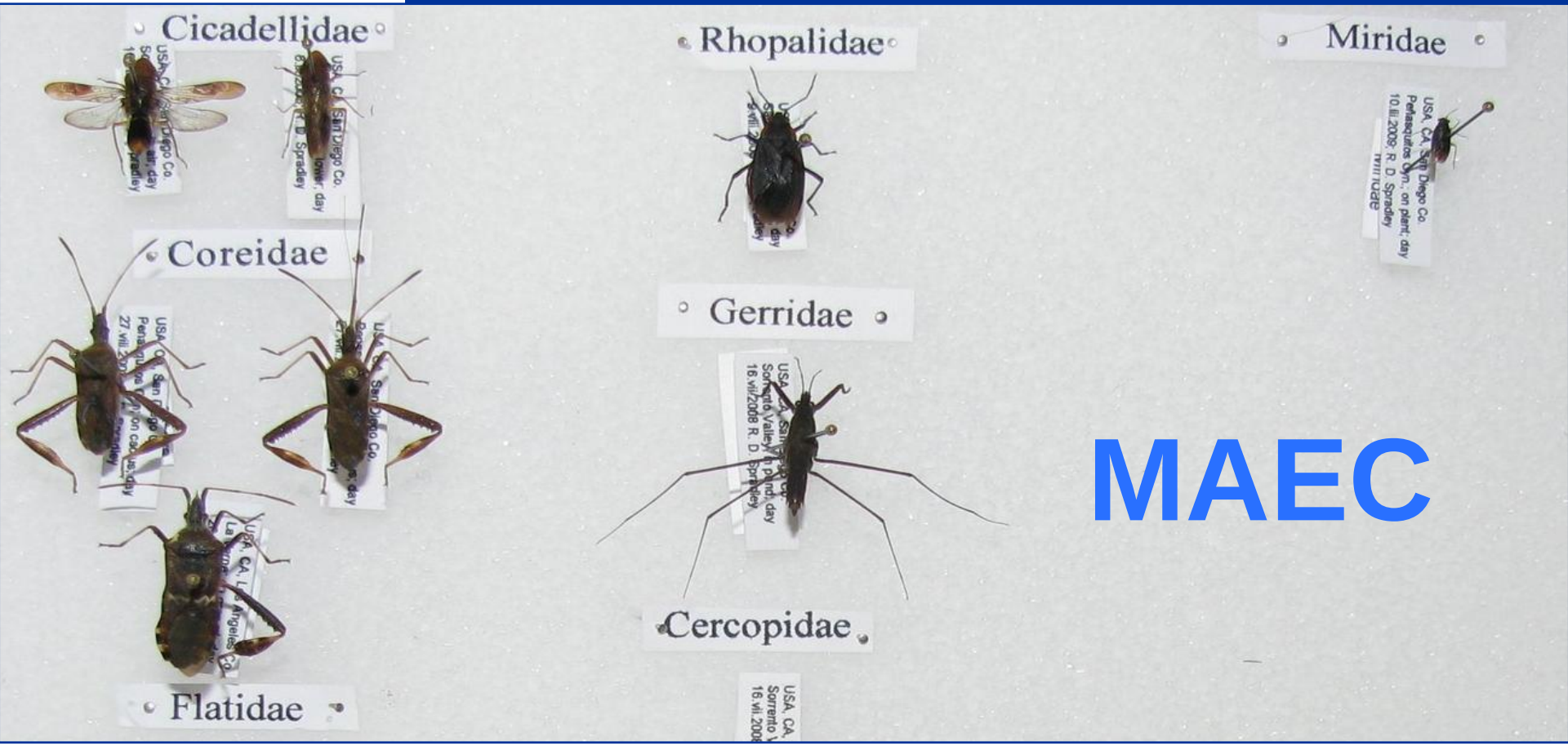


Penny Chase

Ivan Kirillov – Desiree Beck – Robert Martin

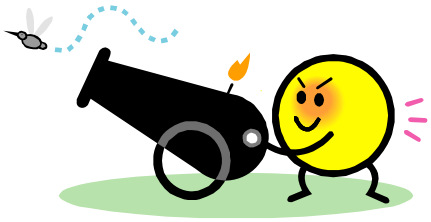
MITRE

Why Do We Need to Develop Standards for Malware?

Multiple layers of protection

Lots of products

Inconsistent reports



There's an arms race

Correlate, Integrate, Automate

Threats

MAEC

Detection

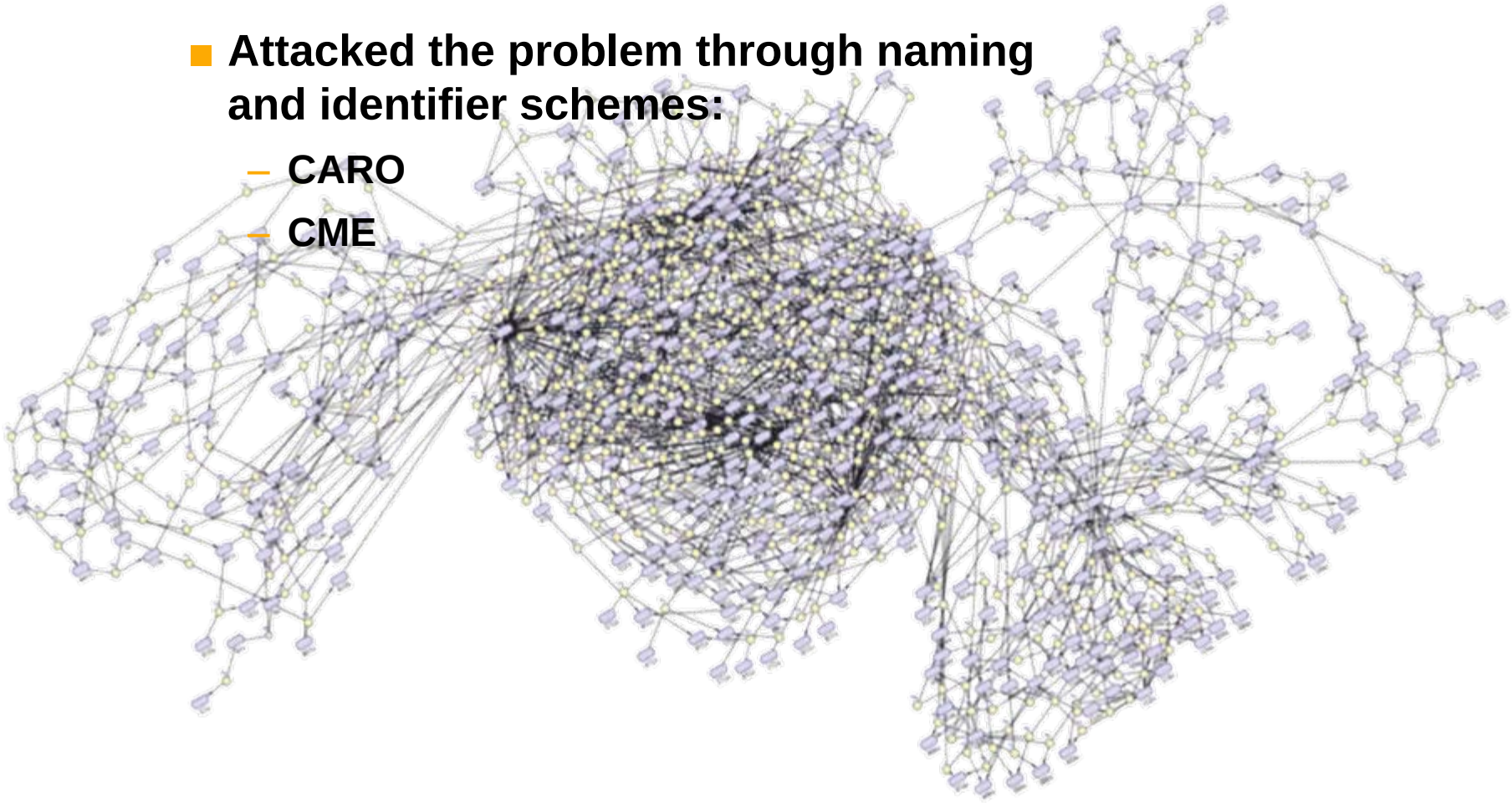
Vulnerabilities

Platforms

Response

Previous Efforts to Bring Order to Malware

- Attacked the problem through naming and identifier schemes:
  - CARO
  - CME



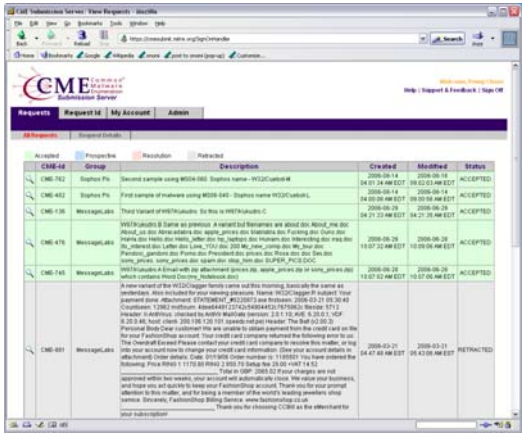
CARO Naming Scheme

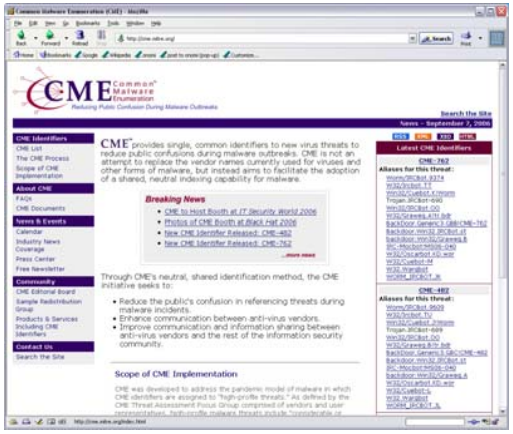
- Created in 1991
- Not an official standard
- Based on encoding attributes as part of the name
  - Type
  - Platform
  - Family
  - Group
  - Length
  - Variant
  - Modifiers
- Vendors have differing implementations
  - W32/MyWife.d@MM!M24
  - W32.Blackmal.E@mm

CME Common™ Malware Enumeration
Reducing Public Confusion During Malware Outbreaks

- **Goals**
  - Decrease public confusion during major malware events
  - Improve communication between antivirus and security products vendors and users
  - Create a neutral, shared referencing capability for malware threats
- **Approach**
  - Unique, common identifier for prevalent malware threats
  - Sample-based process
    - AV Vendors submitted samples
    - CME Board decided when to assign new IDs
    - CME Team created mapping between vendor names and IDs

CME Timeline





Oct 2004

Initial discussions at VB Conference

Nov 2004

Manual test of ID assignment process

Jan 2005

Preliminary Editorial Board established

Feb 2005

Submission Server

Oct 2005

Public announcement and website

Jan 2007

39 CME IDs assigned

CME Community

Editorial Board

- Represent AV industry perspective
- Helped bring the CME concept to maturity

Sample Redistribution Group

- Organizations with access to malware samples
- Submit malware samples for CME identifier assignment

Adopters

- Organizations that use CME identifiers in their products
- Do not have access to the malware sample

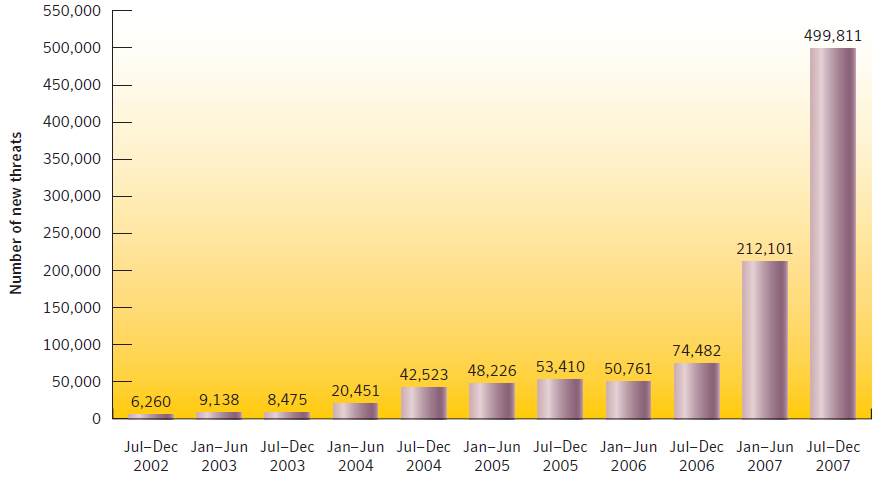
Technical Feedback Group

- Antivirus product user representatives from industry, government , and academia
- Provide user and incident responder needs for CME

The Malware Threat Changed

Rise of New Threats

Symantec Global Internet Security Threat Report, Volume XIII, 4/2008



- **Criminal activity for financial gain remains the driver for the massive increase in Internet threats.**

"Malware Sets Records in 2008", PC World, December 2008

- **Attackers have shifted away from mass distribution of a small number of threats to micro distribution of large families of threats. These new strains of malware consist of millions of distinct threats that mutate as they spread rapidly.**

"Top Security Trends of 2008 and What to Watch for in 2009", Symantec, Information Systems Security

Leading to a Malware Paradigm Shift

- Increased Obfuscation & Armoring
  - Polymorphism
  - Metamorphism
  - Packing
  - Encryption
- Physical signatures falling by the wayside
- New AV Detection Methods
  - Based on heuristics
  - CME's sample-based approach no longer made sense

DHS/DoD/NIST SwA Forum Malware Working Group

- Stood up in 2007 to address concerns of potentially malicious code throughout the system lifecycle
- Goal: Develop a consensus on software that behaves in potentially malicious ways, to
  - Facilitate detection
  - Enable users to make informed decisions
    - Is this software legitimate? Even so, is it appropriate for my environment?
- Co-chairs
  - Ari Schwartz, Anti-Spyware Coalition
  - Penny Chase, MITRE

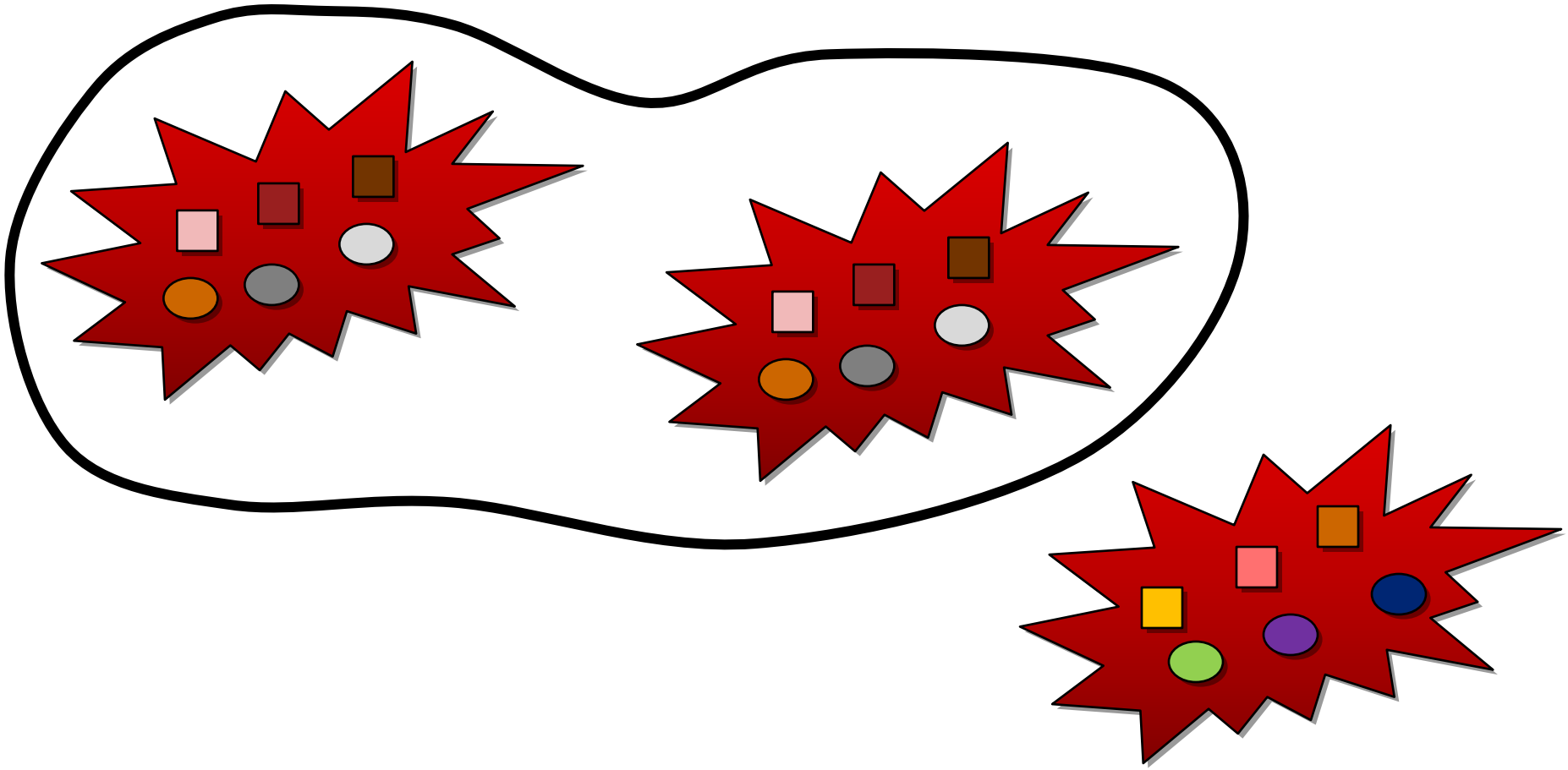
Working Group Approach

- Develop a consistent language of malware attributes and behaviors
- Consider multiple dimensions
- Leverage previous work
  - Anti-Spyware Coalition (ASC)
    - Definitions
    - Risk model
    - Community
  - Common Malware Enumeration (CME)
    - Profile
    - Community
  - Ensure connections with related initiatives
    - CVE, CWE, CAPEC, CEE, OVAL, etc.

Malware Attribute Enumeration and Characterization (MAEC)

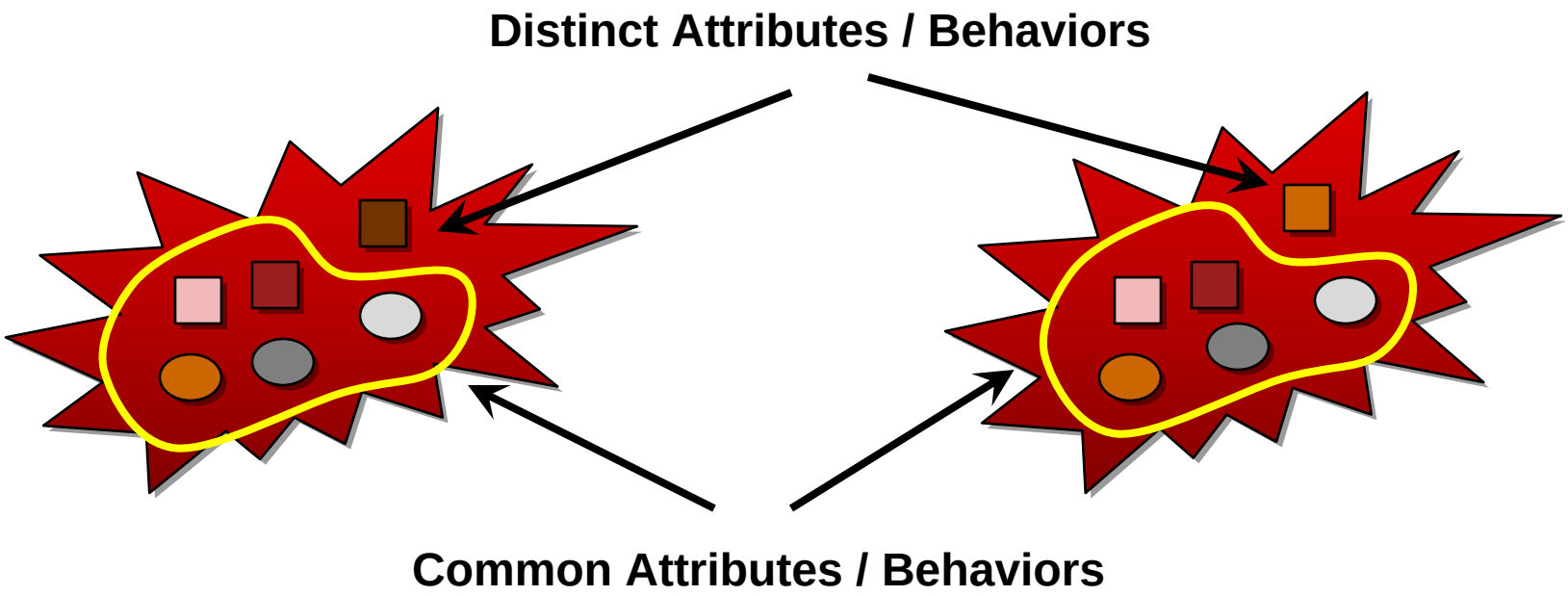
- Formal language for characterizing malware
  - Two core components
    - Enumerated elements (vocabulary)
    - Schema (grammar)
  - Multiple levels of abstraction
- Focus on attributes and behaviors, *not*
  - Intent
  - Malware families

Why Focus on Attributes and Behaviors? (I)



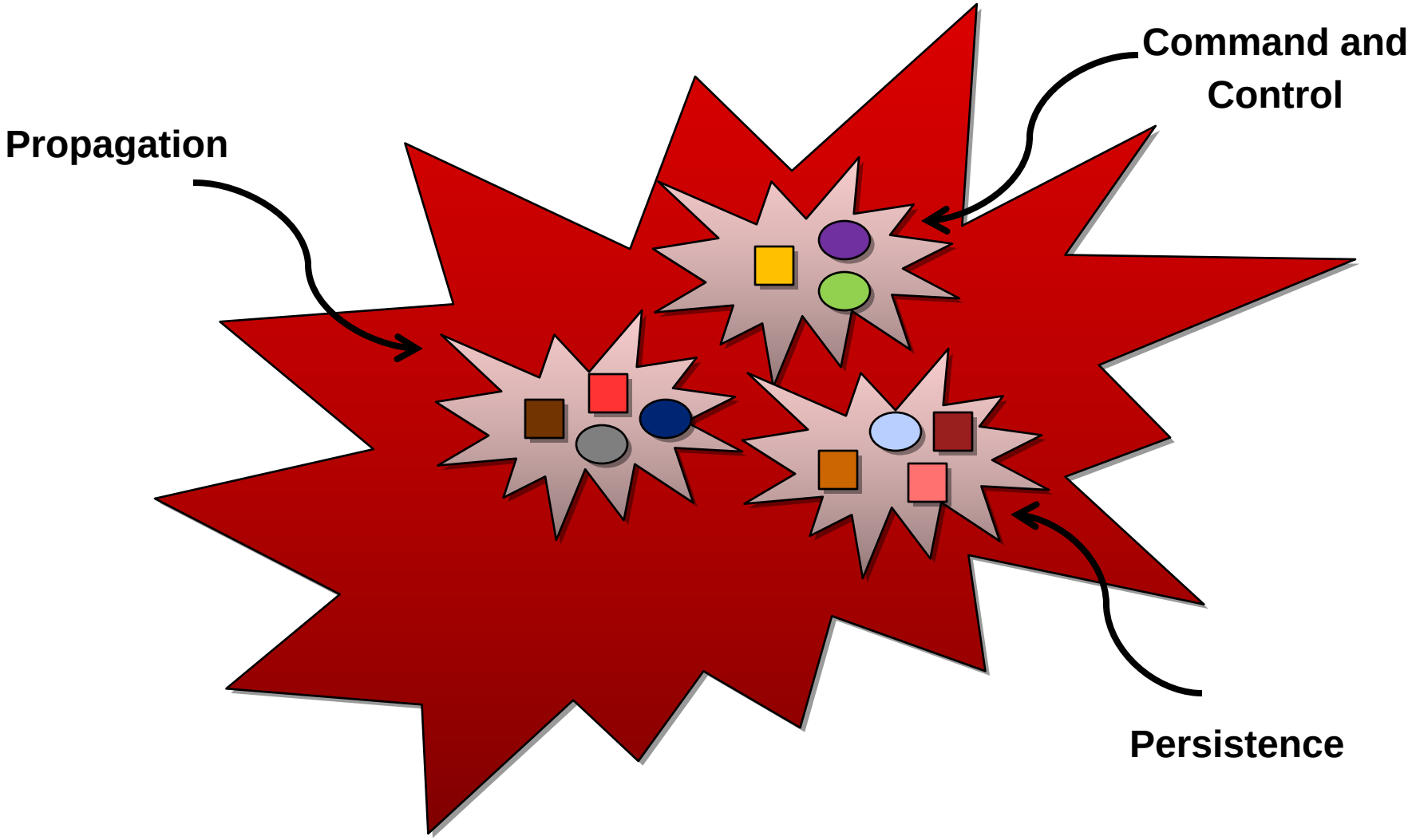
Malware with the same attributes and behaviors can be grouped together

Why Focus on Attributes and Behaviors? (II)



Describe variants in terms of small differences in attributes and behaviors

Why Focus on Attributes and Behaviors? (III)

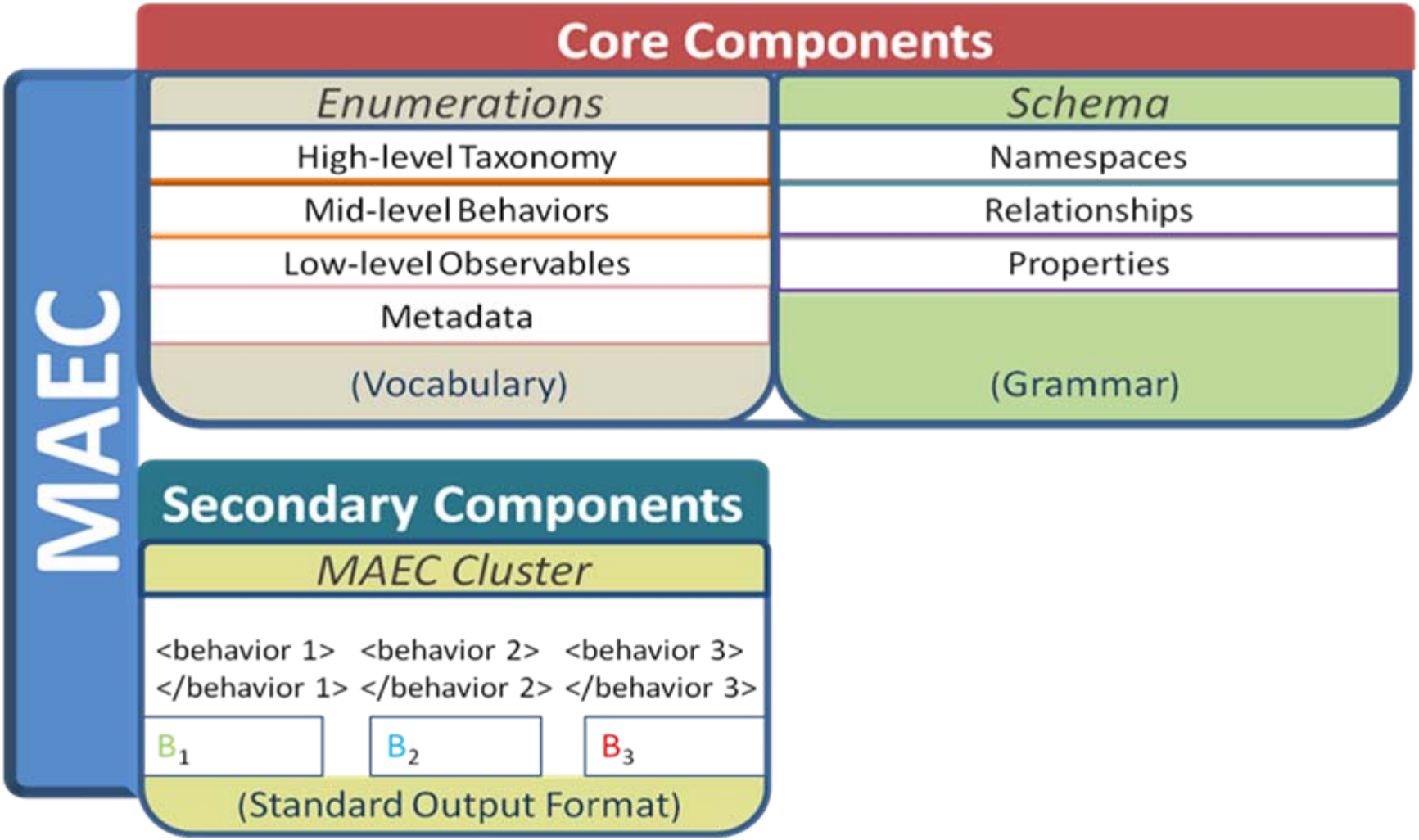


Facilitates describing blended threats

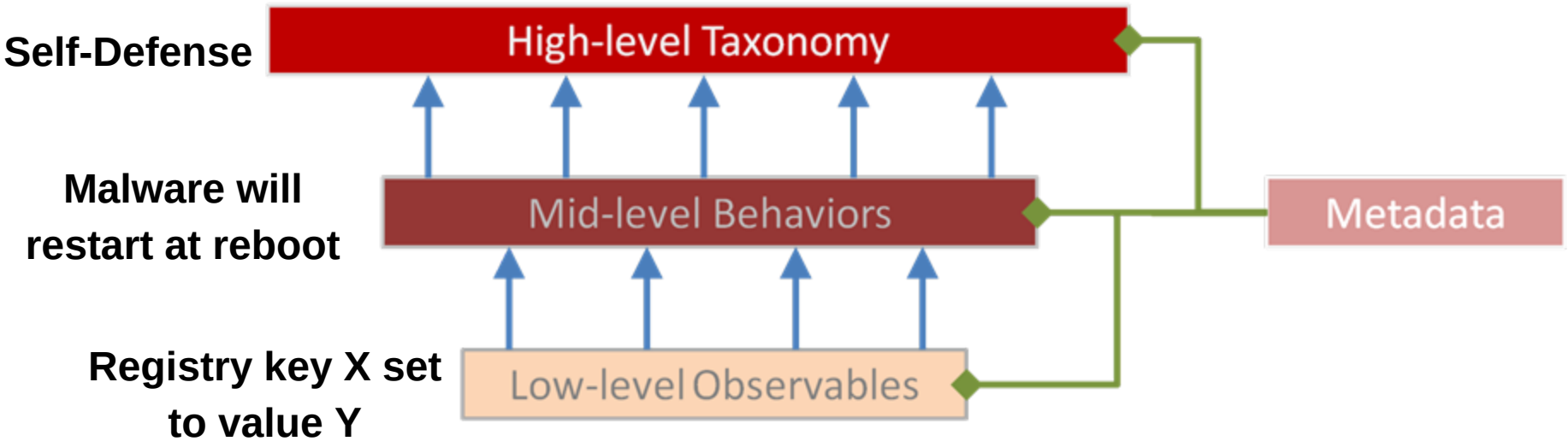
Why Focus on Attributes and Behaviors? (IV)

Mitigates challenges posed by armoring, obfuscation, and polymorphism

MAEC High-level Overview

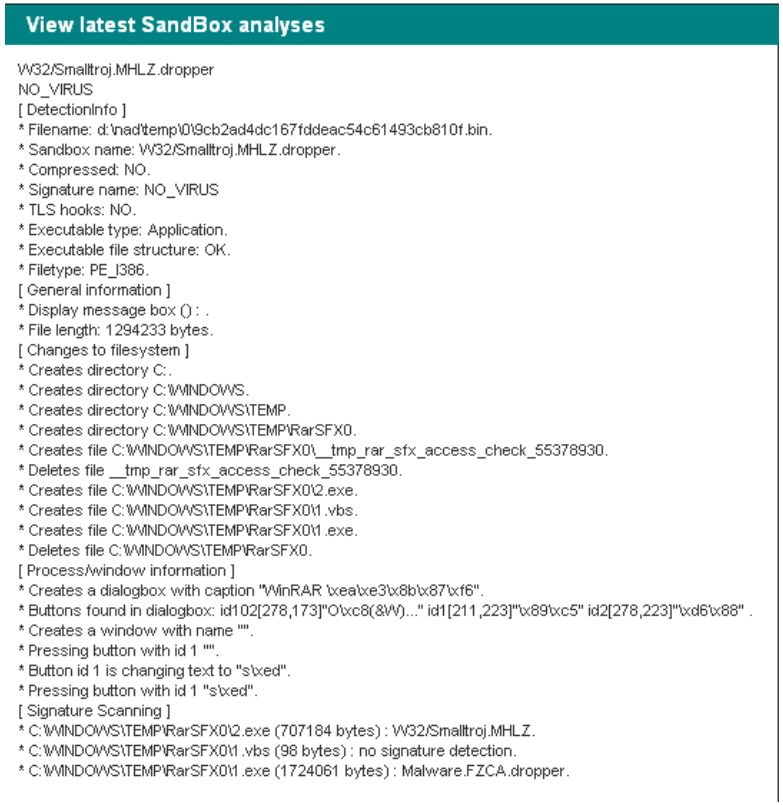


MAEC's Enumerated Elements



Low-level Observables

- System state changes that can pinpoint malware activity
  - Registry Keys
  - Files
  - Network Activity
  - Mutexes
  - Processes
  - Etc.
- Leverage
  - CME Profile
  - IEEE Industry Connections Security Group Malware Group
  - Others ???



Mid-level Behaviors

- Rationales for low-level observables
  - E.g.,
    - What is the purpose of an inserted registry key?
    - Why were several UDP packets sent?
    - Why was this file created?
- Useful for analysis & heuristic detection
  - Links observables to high level behaviors
  - Useful for describing malware authors' TTPs

High-level Taxonomy

- High-level behaviors and characteristics
  - Group together lower level behaviors and attributes
  - Guides analysis
    - Helps ensure there aren't gaps
    - Leads to comprehensive reports
  - Provides multiple views into malware
- Leverage
  - SANS Internet Storm Center Categories of Malware Traits
  - CAPEC
  - ASC
  - Others ???

Metadata

- Relevant non-observable attributes or information about attributes
- Common metadata entities
  - Hashes (MD5, SHA1)
  - Time first observed
  - Etc.
- Attribute-oriented metadata
  - Insertion transparency
  - Dependence of insertion mechanism on user interaction
  - Etc.

MAEC's Schema

- Defines the syntax for the enumeration elements
- Provide an interchange format
- Use standard technologies
  - XML
  - RDF and other Semantic Web technologies?

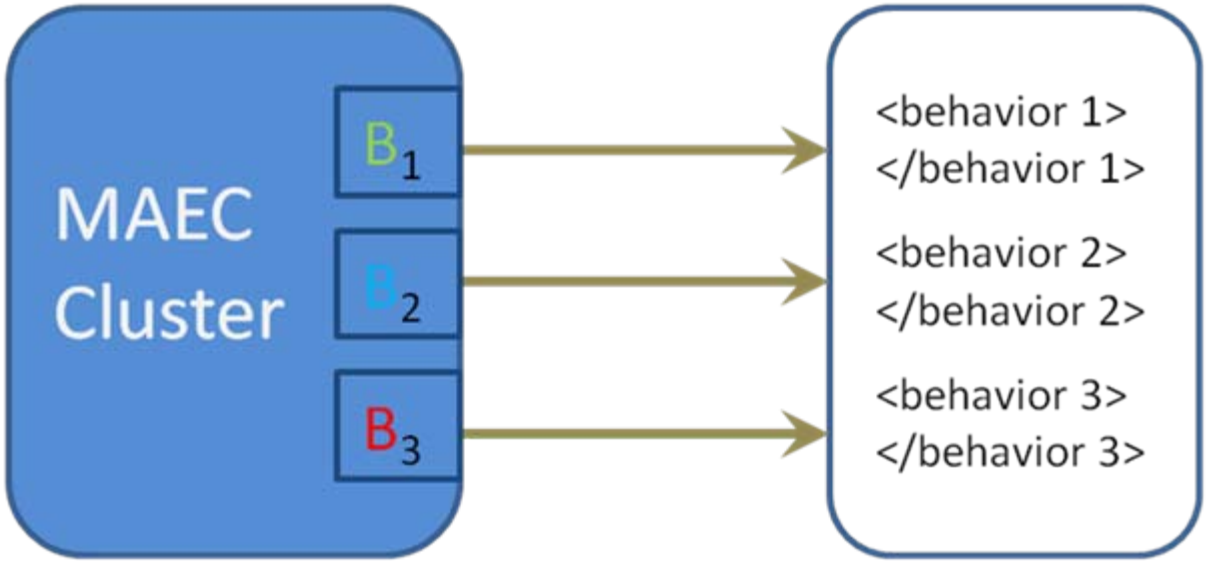
```
<behavior id=1>
    <level>mid</level>
    <type>reconnaissance</type>
    <subtype>keyboardCheck</subtype>
    <attributes>
        <languageChecked>language:ukrainian</languageChecked>
            <successCondition>execution:stop</successCondition>
            <failureCondition>execution:continue</failureCondition>
    </attributes>
</behavior>

<behavior id=2>
    <level>low</level>
    <type>creation</type>
    <subtype>mutex</subtype>
    <attributes>
        <variableStringCreated>"Global\%u-%u"</variableStringCreated>
        <variable>"%u"</variable>
        <variabletype>datatype:decimal</variabletype>
    </attributes>
</behavior>
```

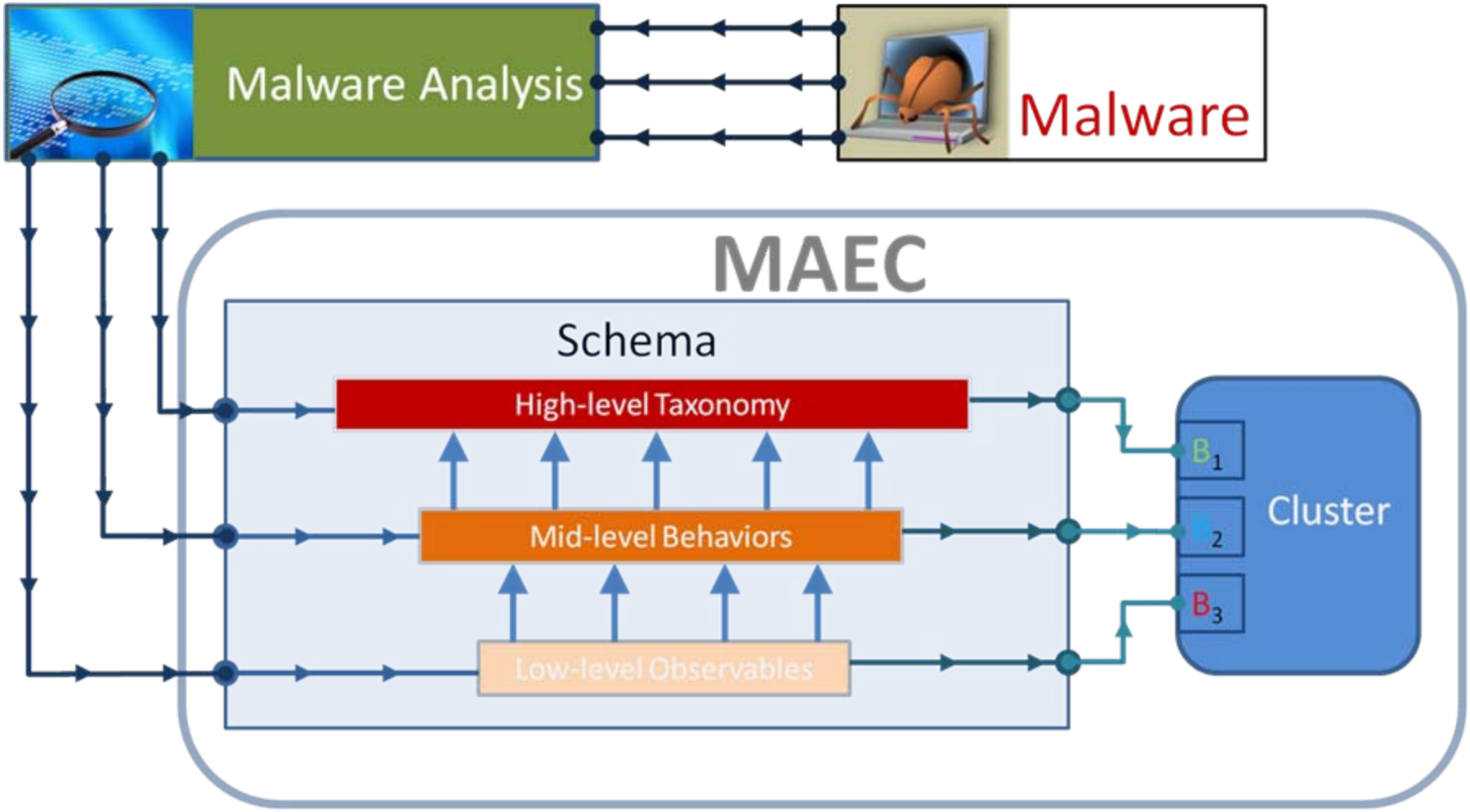
(Sample MAEC Cluster)

The MAEC Cluster

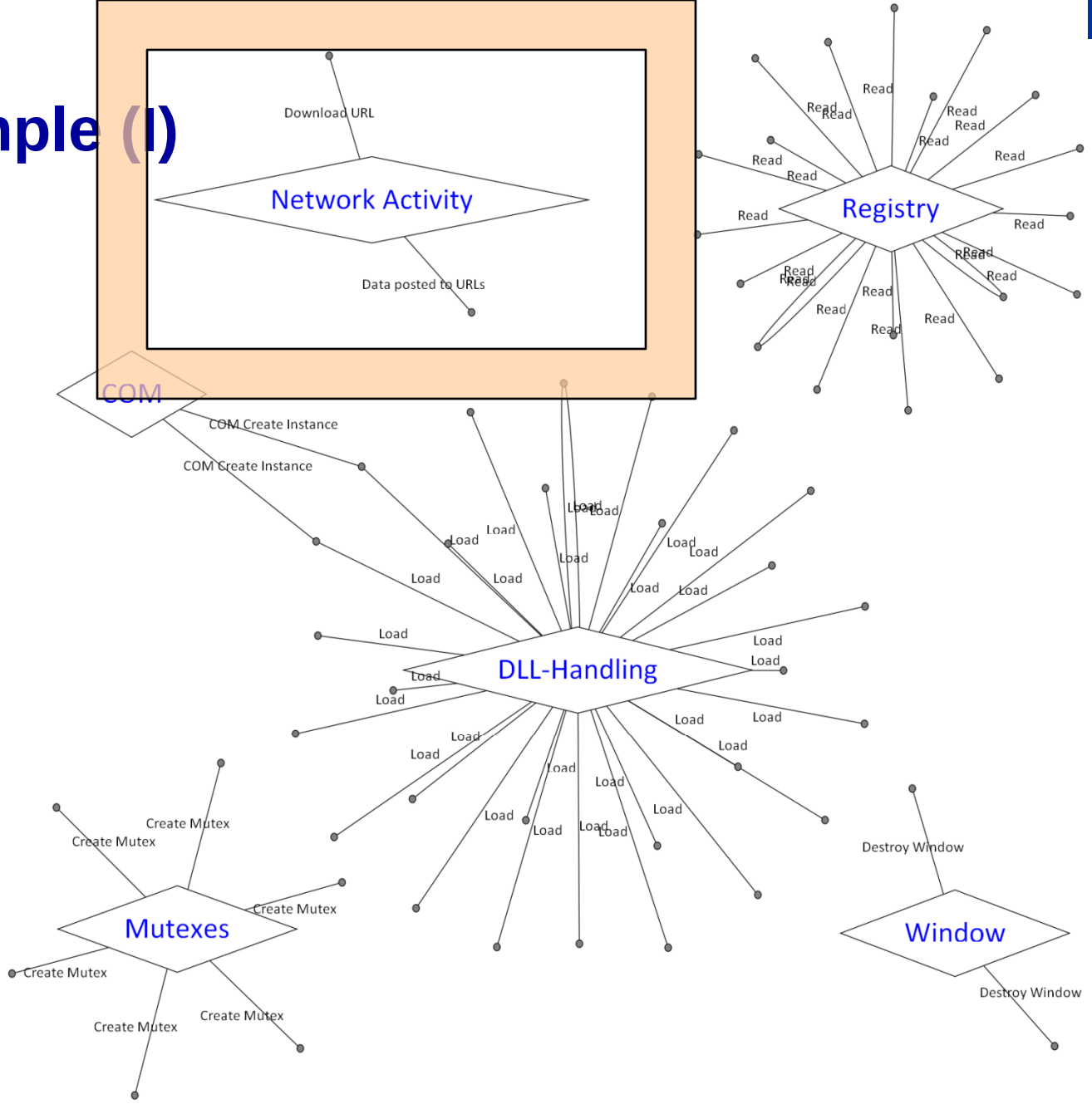
- MAEC-encoded set of attributes for instances of malware
- Standardized form of MAEC output
  - XML
- Can be used to define specific behavioral subsets, etc.



Typical MAEC Usage Scenario

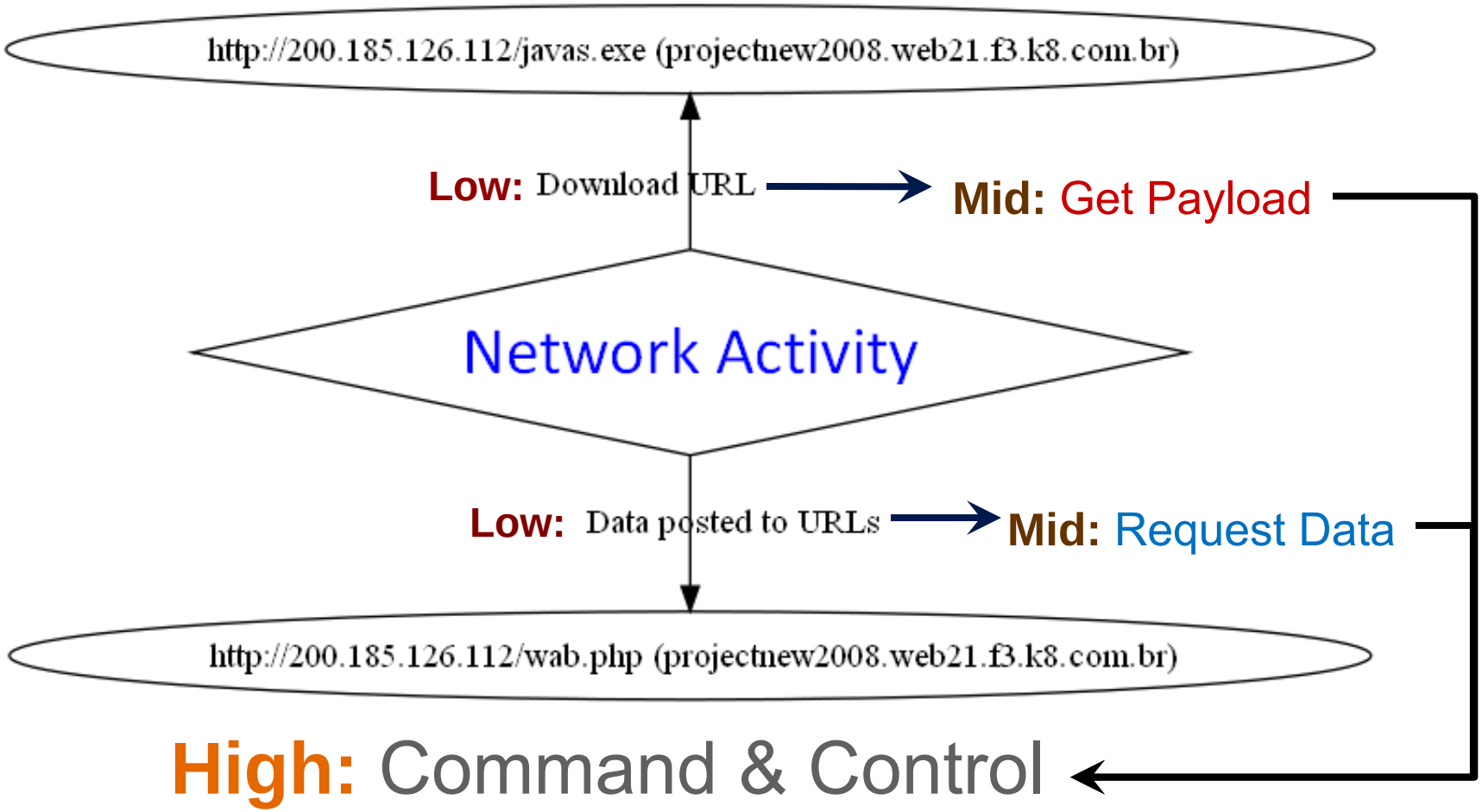


Example (I)

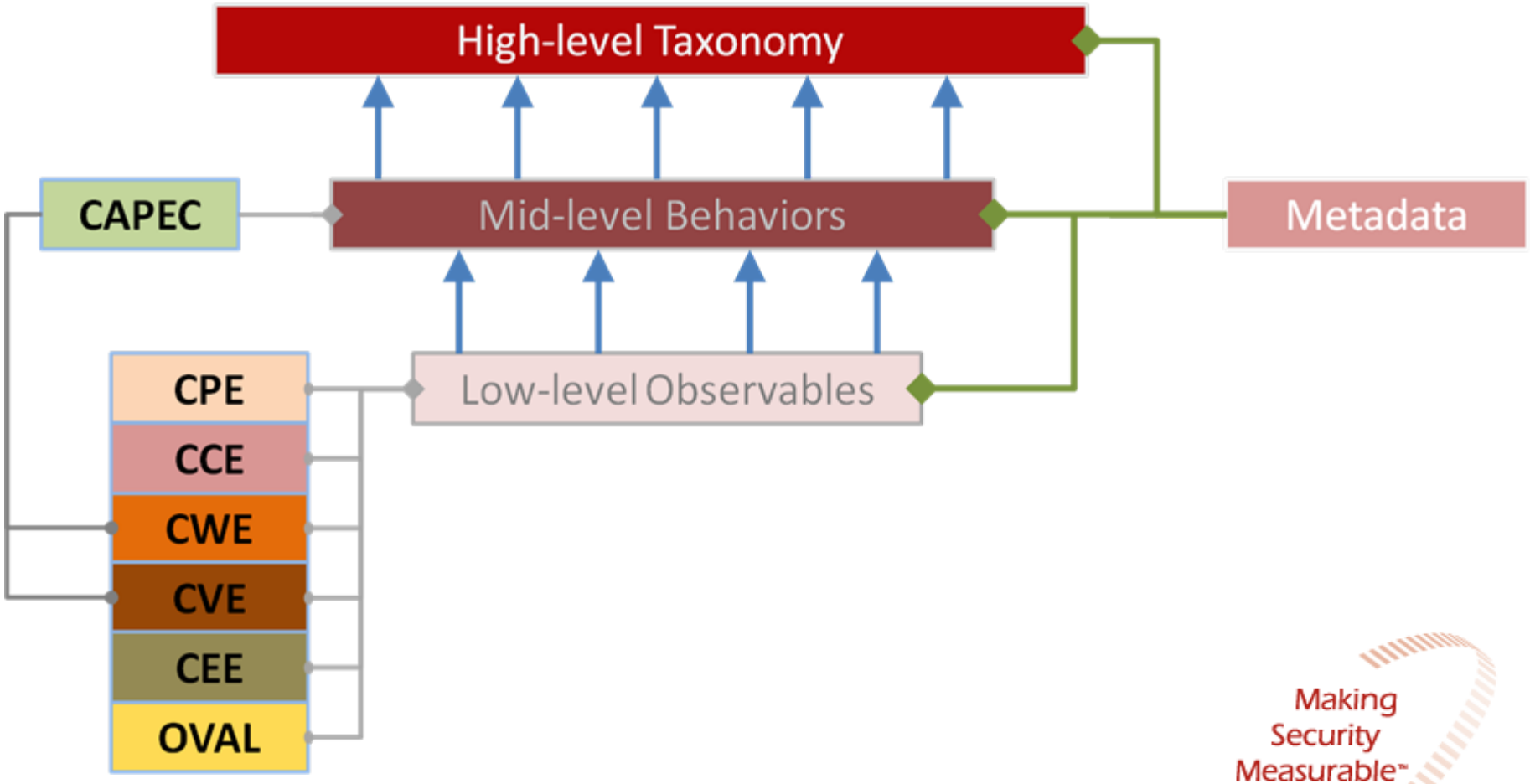


MITRE

Example (II)



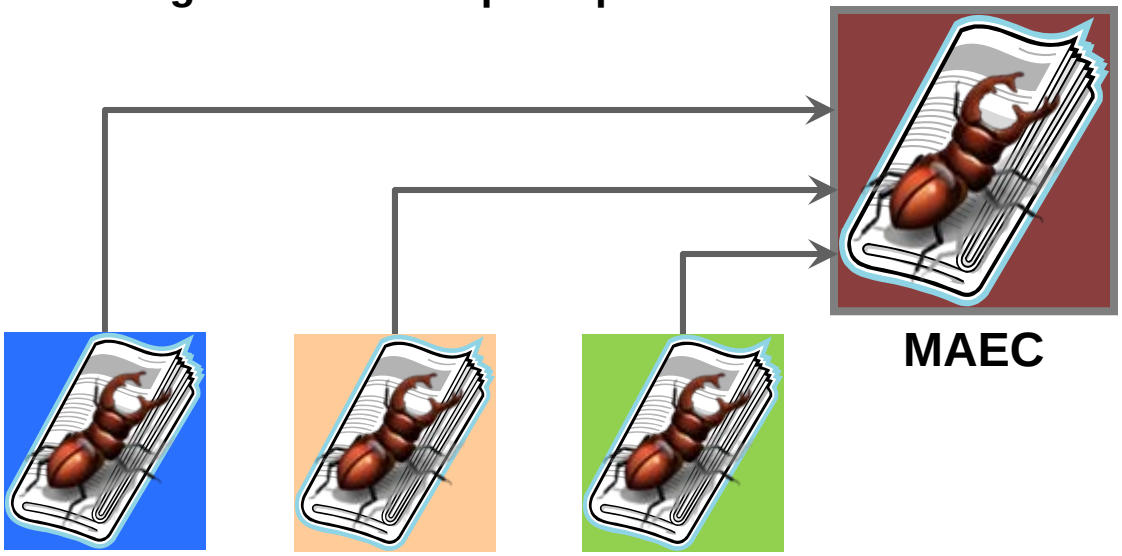
MAEC's Ties to MITRE's MSM Standards



Making Security Measurable™

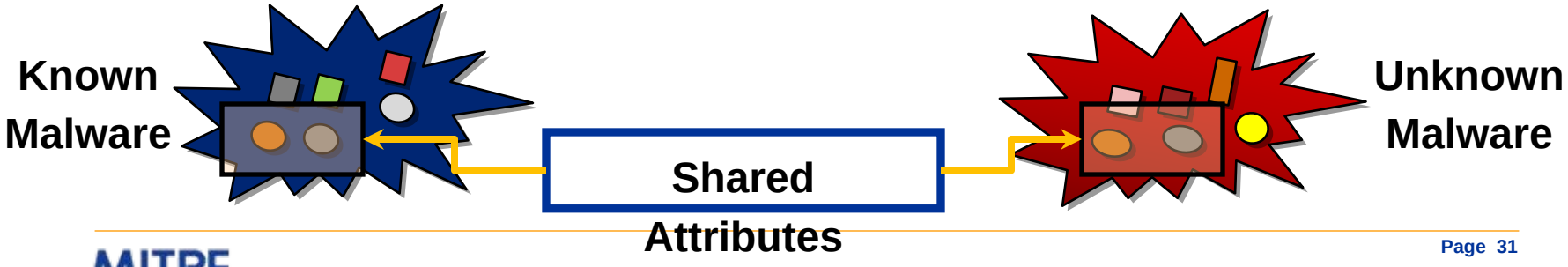
MAEC Use Cases (I)

- Uniform Malware Reporting Format
  - Uniform vocabulary - MAEC Enumerations
  - Uniform structure - MAEC Schema/MAEC Cluster
  - Multiple benefits and uses
    - Reduce ambiguity & confusion
    - Ensure report compatibility
    - Facilitate integration of multiple reports



MAEC Use Cases (II)

- Malware Detection
  - Based on characterized attributes and behaviors
  - Host-based detection
    - Low-level attributes (files, registry entries, etc.)
      - Tool-based detection : OVAL Entries
    - Patterns of mid-level behaviors - heuristics
  - Network-based detection
    - Incoming & outgoing traffic
    - Linkage of traffic to behaviors
  - Detection based on shared attributes
    - Permits detection of new, unknown malware



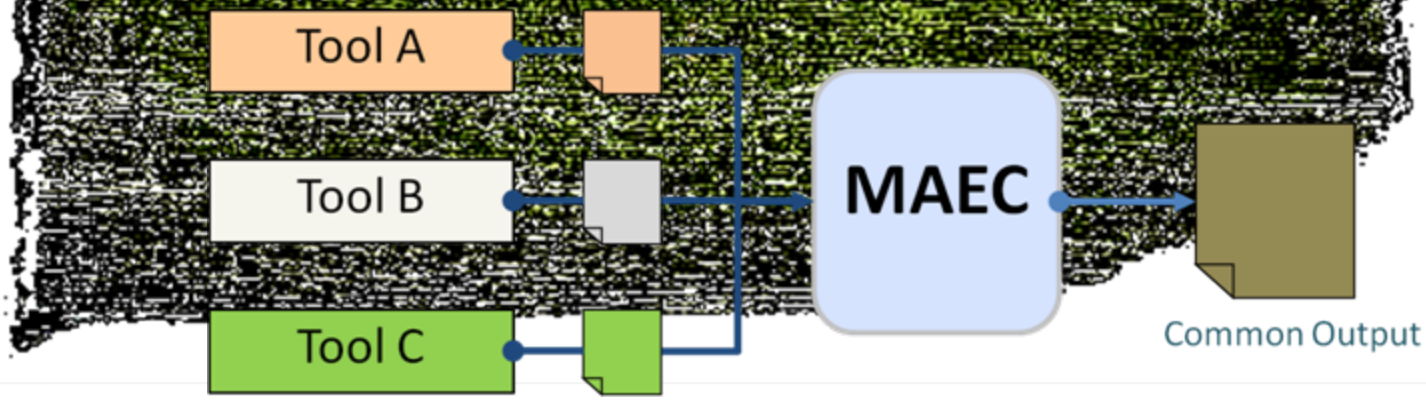
MAEC Use Cases (III)

- **Malware Threat Assessment**
  - **Threat assessment based on MSM links**
    - **CPE – targeted platform**
    - **CVE – targeted vulnerability**
    - **CCE – targeted configuration weakness**
    - **OVAL – check for presence of vulnerability**

  - **MAEC's mid and high-level attributes will provide information on the potential damage caused by malware**
  - **Together, this will facilitate prioritization of malware detection & mitigation efforts**
    - **Malware threat scoring system**

MAEC Use Cases (IV)

- Malware Analysis
  - Standardization of analysis results
    - Common vocabulary – MAEC enumerations
    - No uniform reporting possible without it
    - Facilitate data sharing between researchers
  - Intermediate format for analysis
    - Mapping tool output to common format
      - Behavioral/Dynamic analysis
    - "Tagging" reports with MAEC

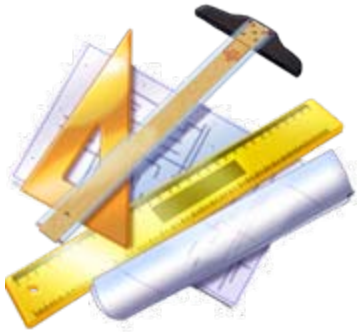


MAEC Use Cases (V)

- Malware Repository
  - Intermediate format for repository schemas
    - Map custom schema to MAEC
    - Facilitate data sharing between repositories
  - Enhanced data-mining
    - MAEC as a repository schema
    - Permits category-based comparison of malware
    - Enables construction of accurate similarity metric

MAEC Use Cases (VI)

- **Objective Criteria for Tool Assessments**
  - **MAEC will define the attributes applicable to specific malware types**
  - **Enables detection tool assessment**
    - **Can the tool detect all possible attributes of the malware types that it claims to detect?**
- **Linking Malware to TTPs**
  - **MAEC provides capability of accurately identifying previously observed tools (malware) used by attackers**
  - **Can be helpful for attribution purposes**
    - **Have I seen these tools used in other attacks? If so, by whom?**



Development Path

- Current status
  - Developing white paper
  - Initial research into MAEC models
    - Focus on low-level observable enumeration
- Standing up infrastructure to support industry, academic, and government collaboration
  - MAEC website
  - Software Assurance Forum working group
- Looking for help
  - Need to build consensus
  - Lots of questions to be resolved
  - Lots of concepts to be fleshed out

Questions?