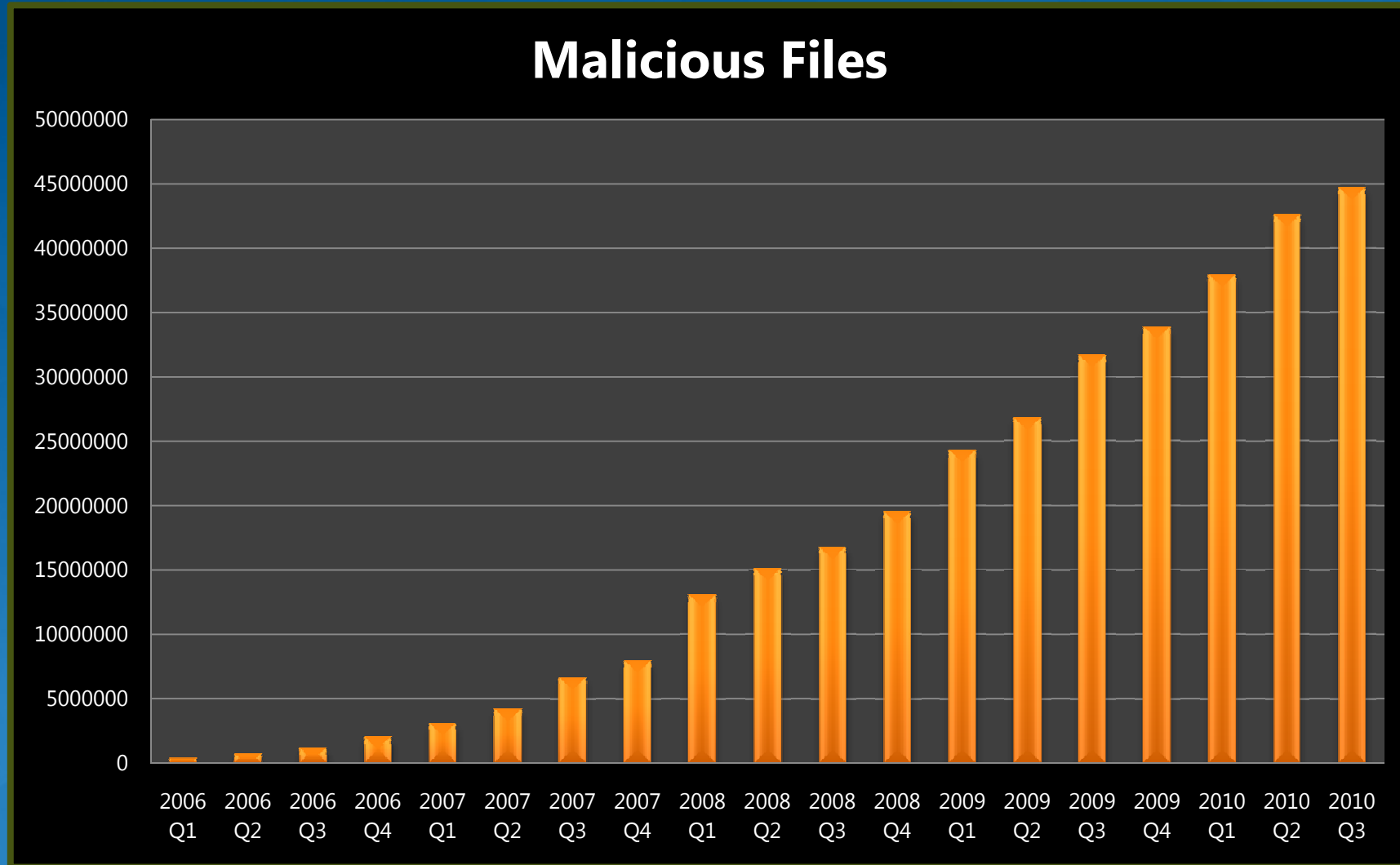


# Telemetry Exchange and Industry Testing

Tony Lee  
Jimmy Kuo

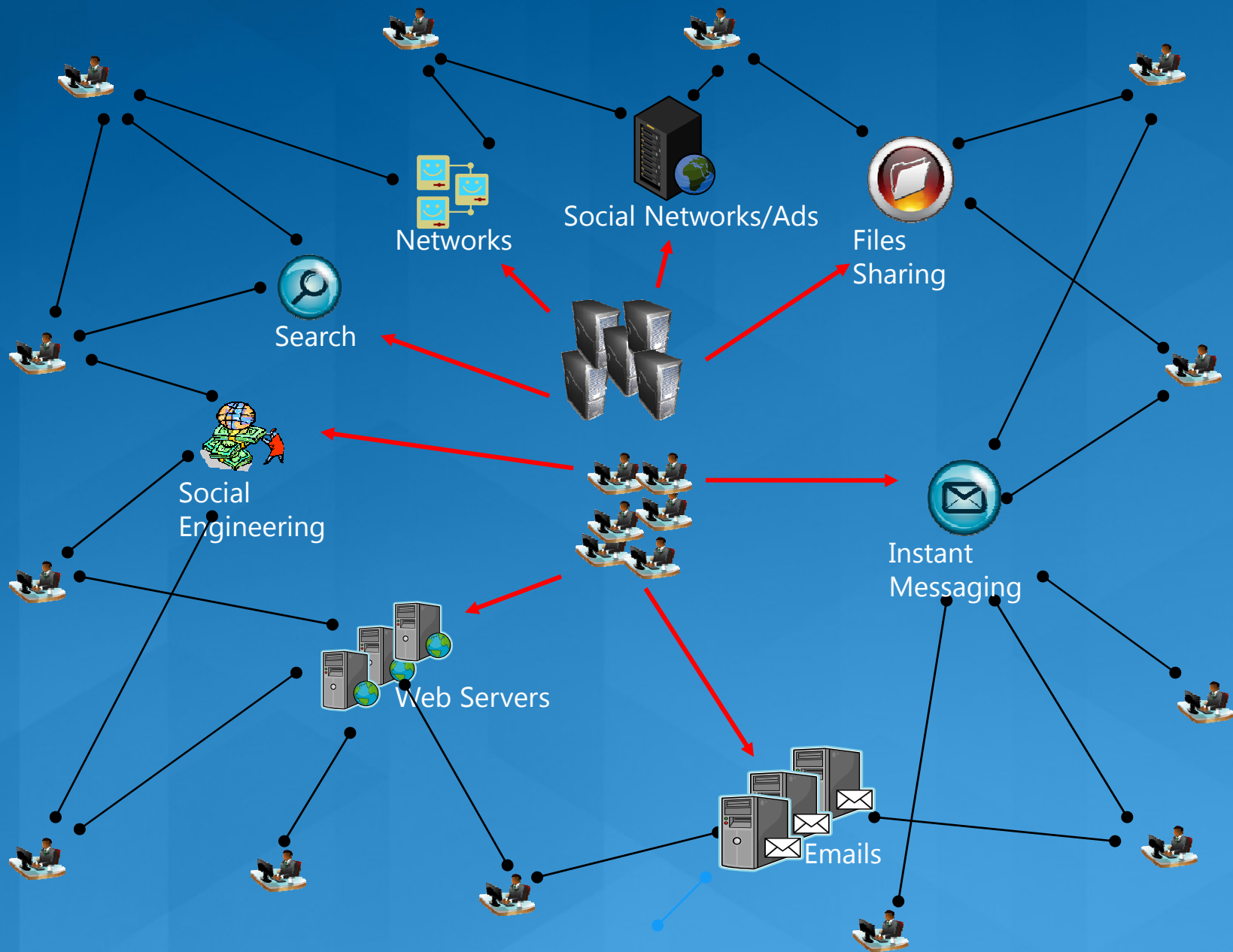
Malware Protection Center  
Microsoft Corporation

# The Number Factor



Source: Microsoft

# The Environment Factor

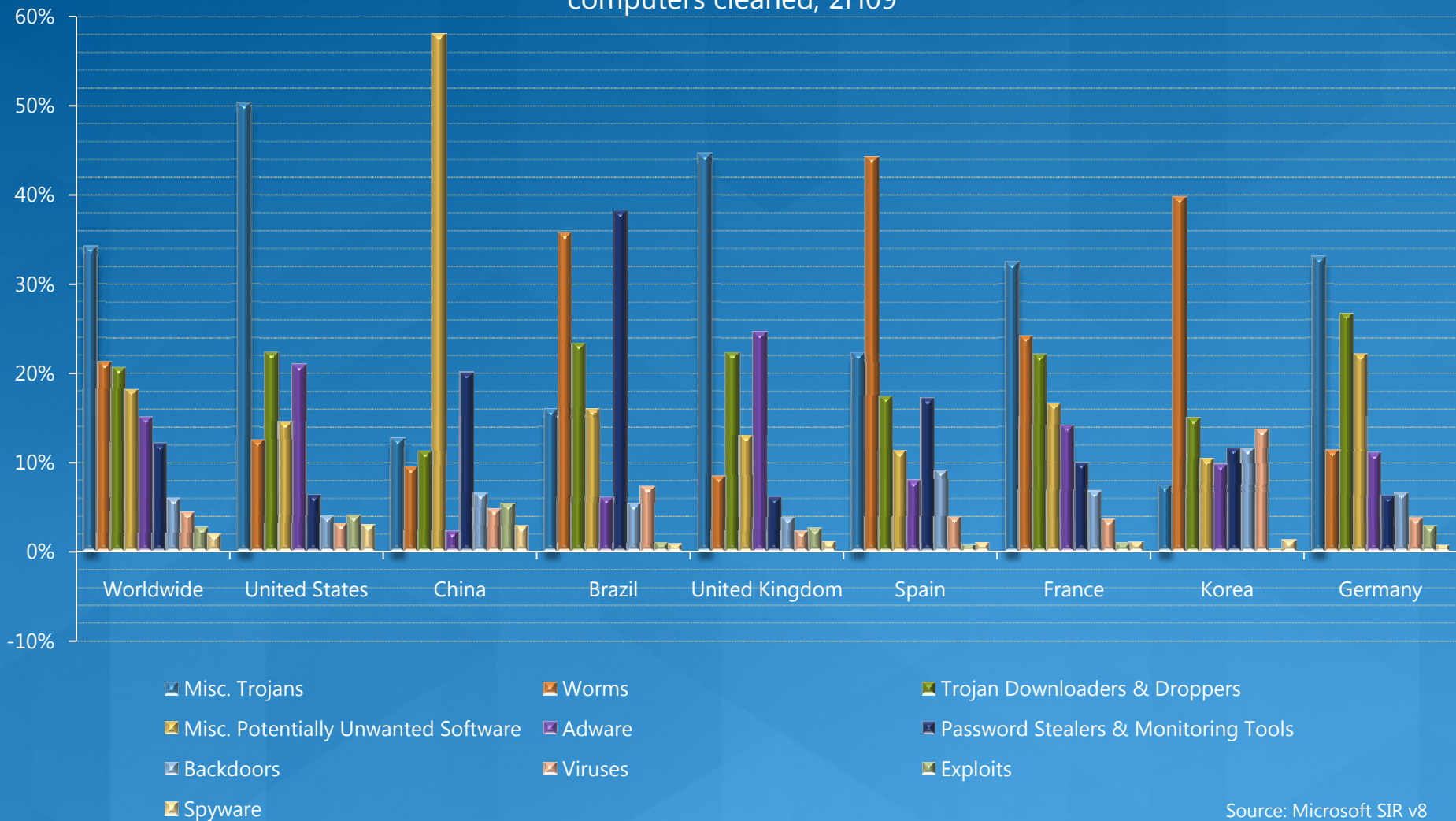


# The Geographic Factor

## Geographic trends

### ● Significant differences in threat patterns worldwide

Threat categories worldwide and in eight locations around the world, by incidence among all computers cleaned, 2H09

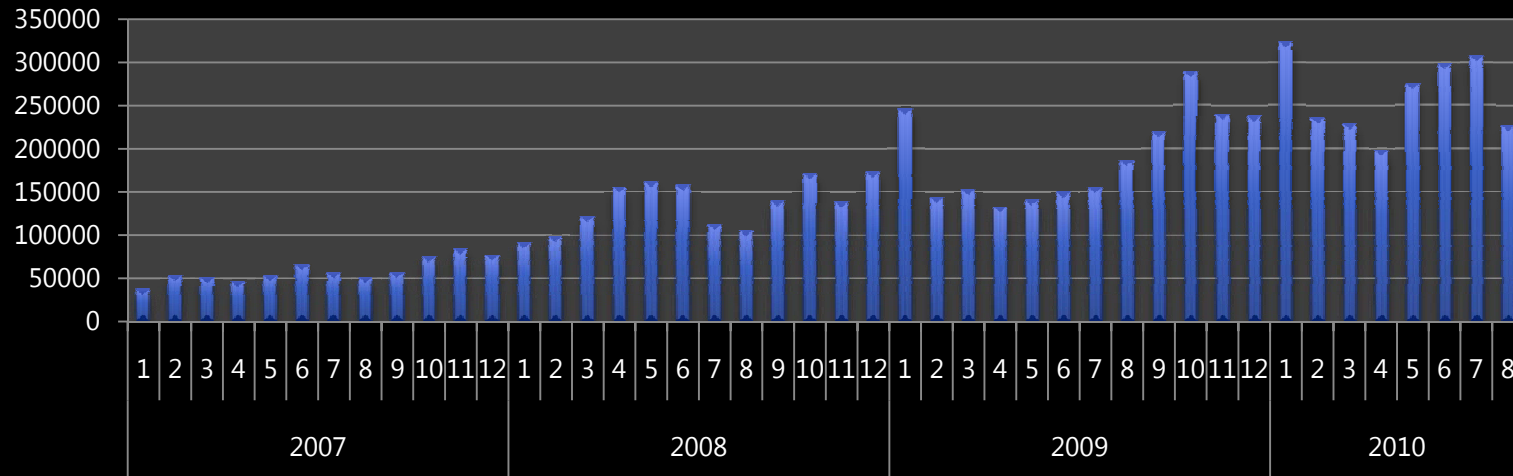


Source: Microsoft SIR v8

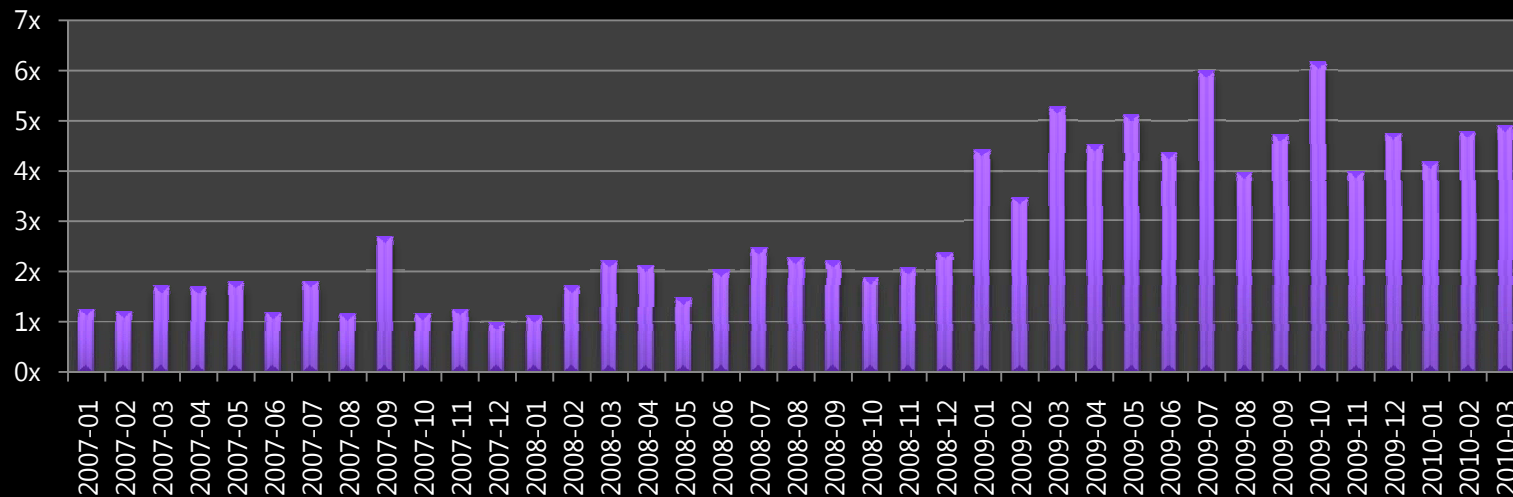


# Threat Cycle Factor

## New Threat Turn-up Trend



## Signature Production Rate



# Complex Threat Landscape

- Large number of threats
- Diverse environment and attack vectors
- Geo-social and linguistic influences
- Shortening of Threat Cycle

## Threat Intelligence Data

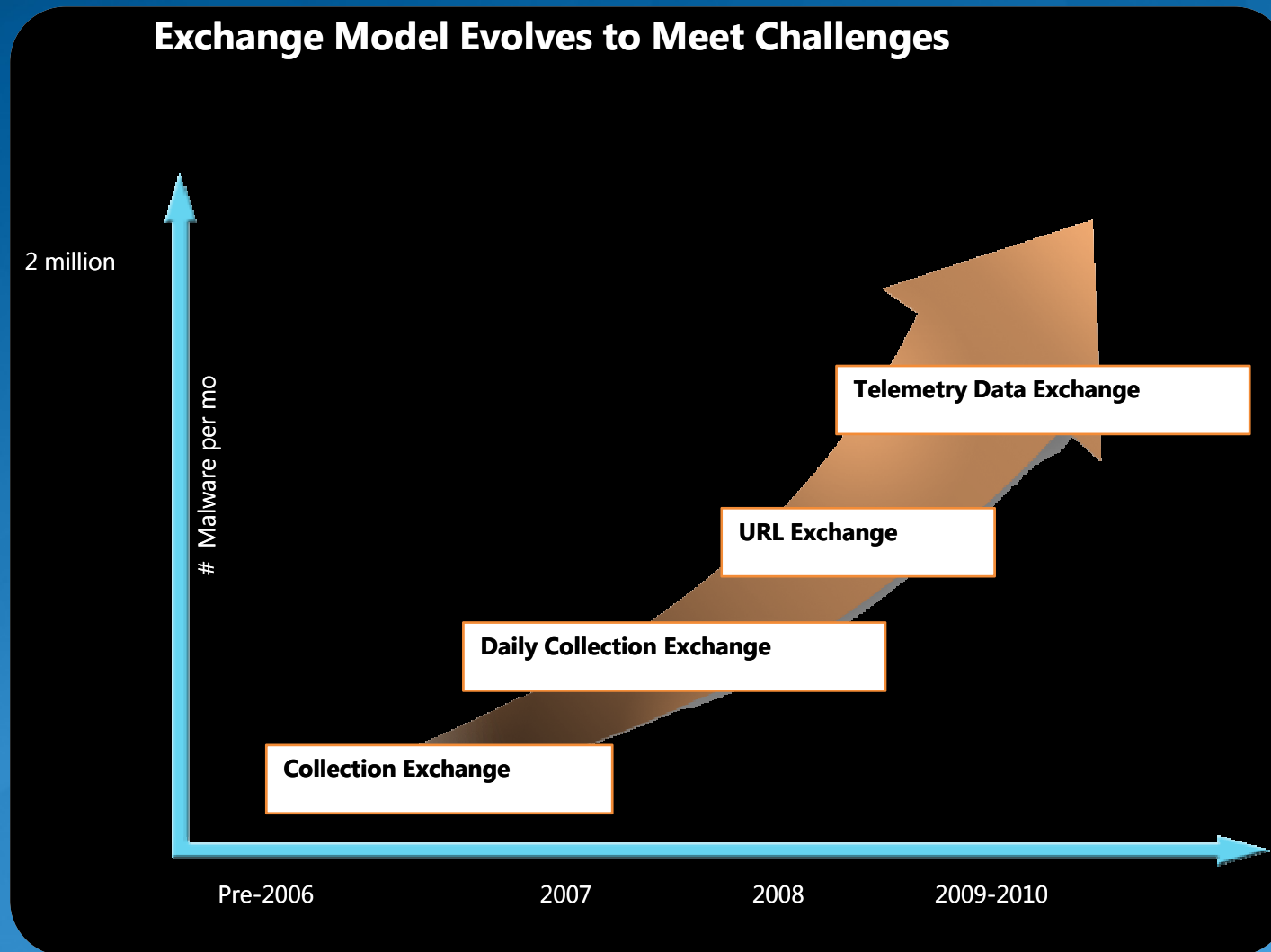


# Information Specialist



**Effective Protection Across On-premise & Cloud**

# Evolution of Collaboration



# Challenges and Solutions

## Challenges

---

## Solutions

Lack of driving  
incentive and  
value

Promote adoption (up the  
stakes in the pot)

Encourage shift in industry  
testing methodology  
toward incorporating  
telemetry data.

---

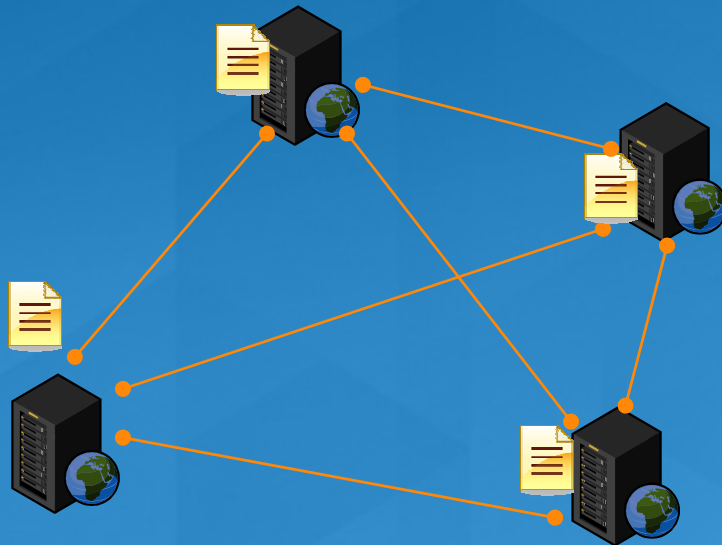
Technical  
Complexity

Educate, demonstrate

Develop and share tools  
and proven methods

# Data Exchanges in Action

- Industry forms working group – IEEE Industry Connection Security Group (ICSG)
- Focused on development of a XML based metadata sharing.
- Exchange are bi-lateral.



AVG  
AV-Test  
Cisco  
McAfee  
Microsoft  
Panda  
SonicWall  
Symantec  
Sophos  
Trend  
WebSense  
and many others...

# Data Exchanges in Action - Schema

<http://grouper.ieee.org/groups/malware/malwg/Schema1.1/>

Detail schema documentation and help

The screenshot displays an XML Schema browser interface. On the left, a sidebar lists various schema components under 'Namespaces' and 'Complex Types'. The main content area shows the 'Component' tab selected, displaying details for the 'classificationDetails' element. The 'Super Types' section indicates a restriction on the 'http://xml/metadataSharing.xsd' namespace. The 'Documentation' section contains a brief description: 'Details of the classification, giving product details, particularly useful for an...'. The 'Properties' section states 'This component is not nillable.'. The 'Model' section shows the XML structure: `<classificationDetails> (definitionVersion?, detectionAddedTimeStamp?, detectionShippedTimeStamp?, product?, productVersion? ) </classificationDetails>`. At the bottom, a 'Nested Element Summary' table lists the 'definitionVersion' element with the type 'xs:string'.

Namespace	Element Name
xs:string	definitionVersion

# Data Exchanges in Action - Scenarios

## Scenarios

Dynamic Prevalence  
Malware File Properties  
Threat Classification  
URL, IP, Domain  
Region  
And many others...

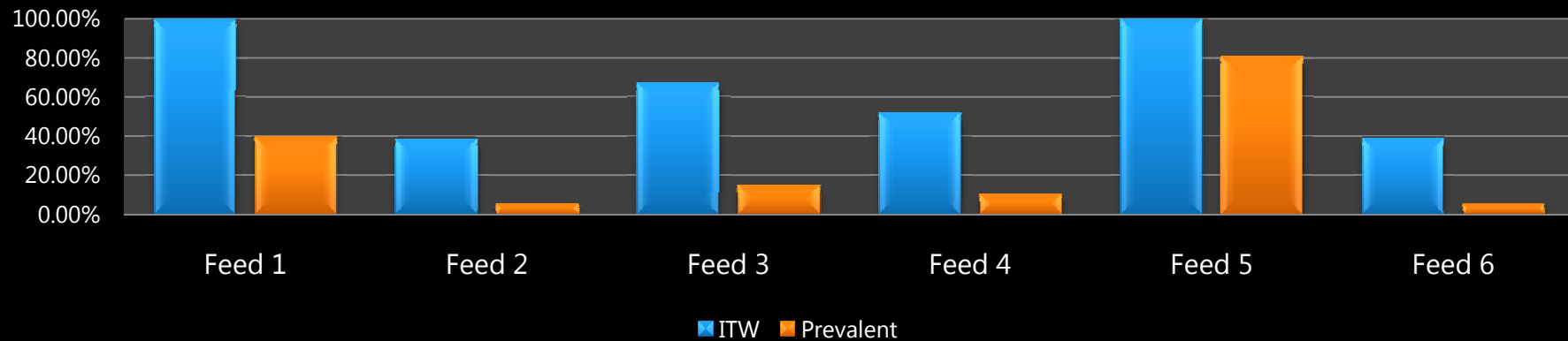
```
<fieldDataEntry>  
  <fieldDataEntry xmlns="http://xml/metadataSharing.xsd">  
    <references>  
      <ref>file[@id="6B8949046F5FD94DF13A0BA386DCAEA4FA1CA6CC"]</ref>  
    </references>  
    <startDate>2009-12-18T04:29:43.150</startDate>  
    <endDate>2010-01-01T04:29:43.150</endDate>  
    <firstSeenDate>2009-07-18T15:33:58.290</firstSeenDate>  
    <origin>user</origin>  
    <commonality>1</commonality> ←  
    <importance>1</importance>  
  </fieldDataEntry>  
</fieldDataEntry>  
<fieldDataEntry xmlns="http://xml/metadataSharing.xsd">  
  <references>  
    <ref>file[@id="D590674EE47A4BBE625E6AEE062D314CA9257D1C"]</ref>  
  </references>  
  <startDate>2009-12-18T04:29:43.150</startDate>  
  <endDate>2010-01-01T04:29:43.150</endDate>  
  <firstSeenDate>2009-07-18T15:33:58.853</firstSeenDate>  
  <origin>user</origin>  
  <commonality>1</commonality>  
  <importance>1</importance>
```

```
</fieldDataEntry><fieldDataEntry>  
  <references><ref>file[@id = 'd79081ae5c380156ce4ab37692083d5e29669aa9']</ref></references>  
  <startDate>2009-10-01T08:54:10</startDate>  
  <endDate>2009-10-28T13:56:31</endDate>  
  <origin>user</origin>  
  <location>DE</location>  
</fieldDataEntry><fieldDataEntry>  
  <references><ref>file[@id = 'd1d7b8e5b1bf31353fea1892e7a49aab59f25da1']</ref></references>  
  <startDate>2009-10-01T10:29:38</startDate>  
  <endDate>2009-10-27T20:56:45</endDate>  
  <origin>user</origin> ←  
  <location>CZ</location>  
</fieldDataEntry><fieldDataEntry>  
  <references><ref>file[@id = '62b90c1f2f66a68b7e40527ca27c461b09e4bda2']</ref></references>  
  <startDate>2009-10-01T00:31:44</startDate>  
  <endDate>2009-10-28T15:06:22</endDate>  
  <origin>user</origin>  
  <location>KR</location>  
</fieldDataEntry><fieldDataEntry>
```

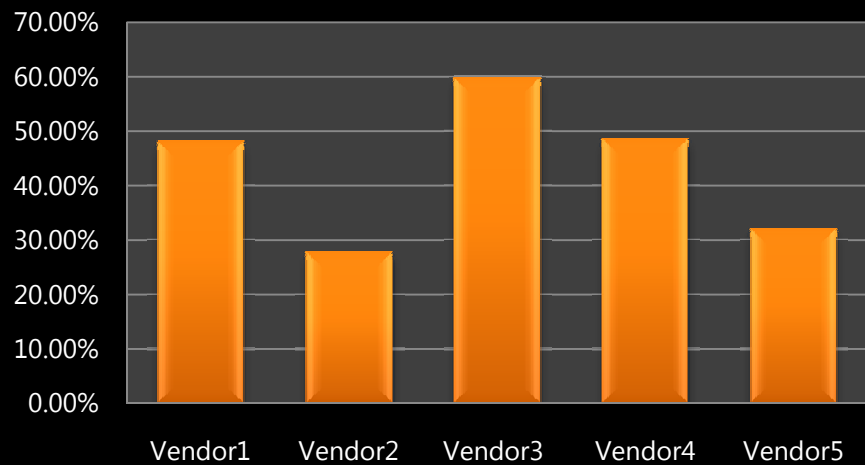


# Data Exchange in Action – Case Study

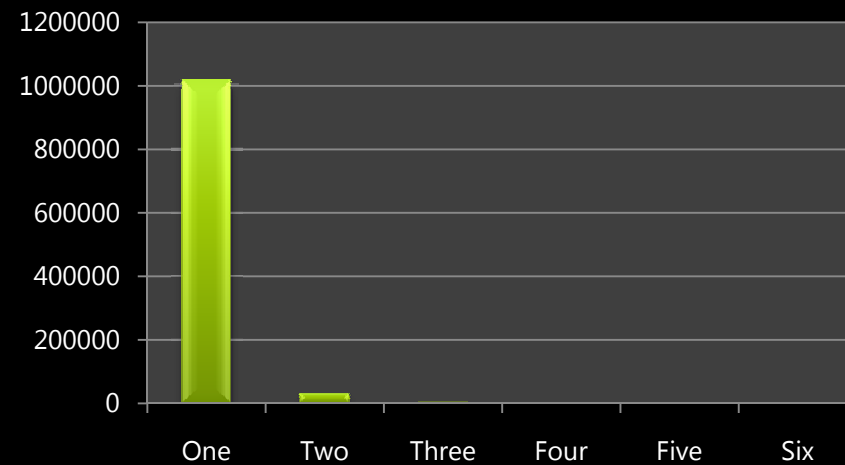
## ITW Quality



## Detection Rate on New Threats

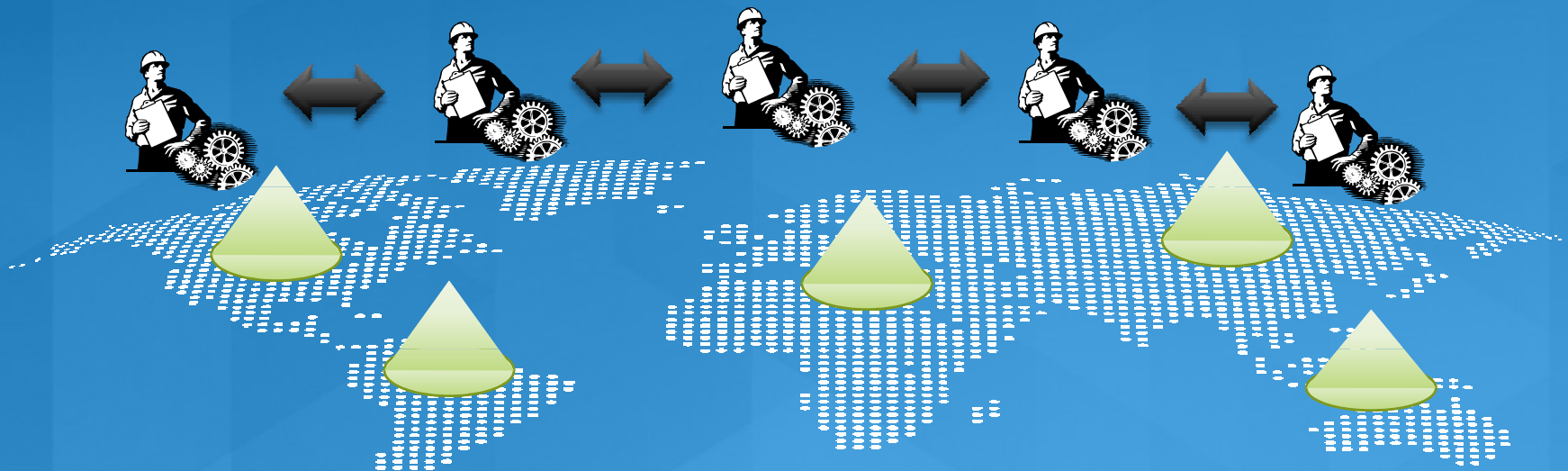


## Common Sources



# Data Exchange In Action

- Data feeds show promising value.
- Standard XML schema reduces cost/complexity, improve data representation.
- Share tools and experience



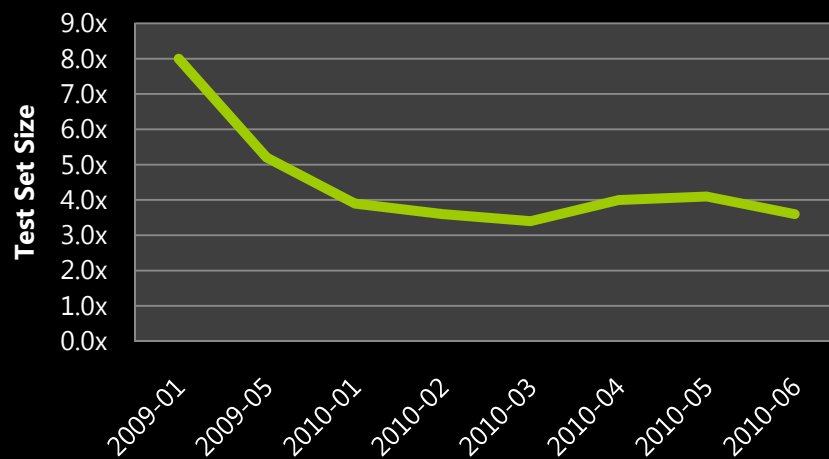
# Industry Testing

## Relevance Measure in Testing Methodology

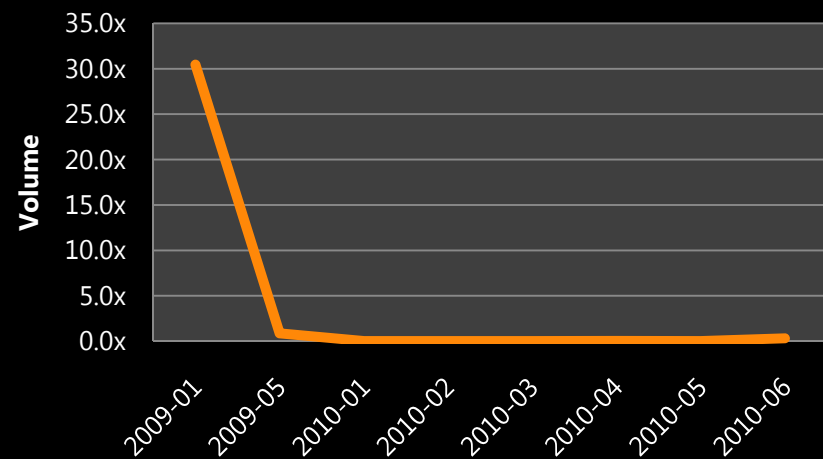
- More samples  $\neq$  Better test
- Be representative of what users really see
- Reflect product capability in user protection

Testers made substantial progress in refining test set quality, leveraging data feeds and tools from industry.

### Trend of Test Set Sizes



### Invalid Files in Test Sets



# Industry Testing Cont.

## Concerns

---

Participants gaming test with faulty data

---

Complexity in merging disparate data sources

---

Lack of data sources

Industry receptiveness to new tests

## Solutions

---

Collate data from different sources, prevent subset of feeds from dominating, not much different from samples.

---

IEEE ICSG Data Exchange Schema  
Source normalization

---

Need to take a step forward to set in positive feedback cycle.

Add additional sub-test or create experimental test, for example, test with telemetry-base score.

Work with IEEE ICSG, industry wants and welcomes changes.

# What do these all mean to users

- AV products can better protect customers, benefited from threat intelligence sharing.
  - Signatures on “real” threats can go into the delivered sets. Less than real threats can have their signatures in the cloud.
- Product testing results are more indicative of protection performance in the field, i.e. what impacts the actual users.
- As test sets get smaller (full product tests), telemetry based sets become even more important.

# The Importance of False Positives in Testing

# How good is The Perfect AntiVirus?

- 100% detection!
- Never needs updating!
- VERY FAST!
- Low System Overhead!

# Hypothetical Test Results

AV-Comparatives On-Demand Test Aug 2010: Tested

AV-Comparatives Proactive Test May: Advanced

AV-Comparatives Whole Product Test: Advanced+

AV-Test: 100% Protection and receive certification

PCSL Total Protection Test: 97% (if no files in clean set)

NSS: 100% detection of social engineering malware



# Hypothetical Certification Results

ICSA: Fail!

West Coast Lab: Fail!

VB100: "20 years of trying, the VB100 continues to be out of its grasp. I don't know why it continues to get satisfactory marks from testers, nor even why I continue to test it! I suppose it's because it's so easy to test. You can always stand by this product to say the same thing time after time." -- John Hawes

# The Perfect Antivirus

Kept as a secret from consumers for 20 years:

v1: echo "%1 is infected."

Ask someone for v2! It's even better. No FPs!

# Conclusion

If The Perfect AntiVirus does not fail the test, the test is as silly as this example.

Don't forget about v2.

# *Microsoft*<sup>®</sup>

© 2010 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.