Functional Specifications

• Detect and cluster malware
• Provide a Web Dashboard for analysts
• Provide a framework for assigning signatures to new malware
Design Specifications

- Malware analysis tool
- Malware aggregator
- Nodal graph display
- Malware statistics applets
Screen Mockup: Nodal Graph

Batch Information
Files Processed: 103/250
Time Elapsed: 00:15:23
Flagged Files: 15

Flagged Files
TMP1CE2.TMP
Memo.pdf
DriverPack.exe
Virus.exe
NotSpyware.jar
Screen Mockup: Flagged Filtering

Batch Information
- Files Processed: 103/250
- Time Elapsed: 00:15:23
- Flagged Files: 15

Cluster Analysis
- Total Number of Files: 15
- Behavioral Profiles:
  - 4 files behave like Trojans
  - 6 files behave like Viruses
  - 3 files behave like Rootkits
  - 2 files behave like Worms

Flagged Files
- TMP1CE2.TMP
- Memo.pdf
- DriverPack.exe
- Virus.exe
- NotSpyware.jar
Screen Mockup: File analysis

**Batch Information**
- Files Processed: 103/250
- Time Elapsed: 00:15:23
- Flagged Files: 15

**File Analysis**
- **File Name:** WannaCry.exe
- **File Signature:** File matched known signature of Ransom.Wannacry
- **Behavior Profile:** Behaves like a trojan and a worm
- **File Size:** 5.65MB

**Flagged Files**
- TMP1CE2.TMP
- Memo.pdf
- DriverPack.exe
- Virus.exe
- NotSpyware.jar
Technical Specifications

- Static analysis module
- Dynamic analysis decision logic
- Malware sandboxing
- Signature information framework
- Malware aggregator
- Database
- Web Front End
System Architecture
System Components

• Hardware Platforms
  ▪ Proofpoint hardware cluster
  ▪ ESXi HyperVisor
  ▪ Linux Ubuntu VMs

• Software Platforms / Technologies
  ▪ Python, Yara, Cuckoo, ClamAV
  ▪ SQLite, Apache
  ▪ Postman API and Bootstrap Library
Risks

• Malware Clustering and Categorization
  ▪ Clustering malware based on file characteristics
  ▪ Research the best way to cluster malware (PE Hash or Fuzzy hashing)

• Understanding Dynamic and Static Analysis Tools
  ▪ The tools behave differently and output different formats
  ▪ Running different malware samples and analyzing outputs

• Scalability and Speed
  ▪ Analyzing variable amounts of malware in an efficient way
  ▪ Properly allocate resources for

• Signature Generation Framework
  ▪ Provide a way for analysts to easily create the signature of a malware
  ▪ Determine what analysis information is relevant for a signature
Questions?