2022年CCF-深信服伏羲基金申报课题

课题一：低编码计算量的视频编解码方案

研究背景：虚拟桌面（VDI）是满足单位办公场景下数据可控和统一运维需求的主流技术方案。深信服公司是国内最大的VDI软硬件厂商，产品获大量机关企事业单位采用。VDI可理解为在后台服务器上运行大量的 Windows/Linux 虚拟机，并通过桌面传输协议，与客户端（硬件终端或软件终端）连接，从而实现的远程桌面。由于涉及大量桌面图像数据的传输，为提升网络性能，改善用户体验，需要在服务器端对桌面图像进行编码，并在客户端解码，以减小网络传输的带宽压力。然而，目前的视频编解码方案如 H264/H265 均为编码复杂而解码简单。复杂的编码过程对服务器造成很大计算压力，而客户端的设备（尤其是笔记本和手机等智能设备）拥有的计算能力又被大量浪费。

研究内容：

1. 研究并设计视频编解码方案，降低编码计算量，实现编解码计算量从服务器端向客户端转移，缓解服务器压力的同时充分利用客户端计算资源。

课题提供：

视频传输测试环境，包括服务器和客户端硬件设备、软件调用编解码算法接口、测试用的视频、原始视频与编码后视频的质量对比评价方案、用于测试对比的h264编解码方案

考核指标：

1. 视频压缩率及画质与对比基线（课题组提供的h264编解码器）持平
2. 编码的CPU消耗比基线降低30%
3. 解码的CPU消耗不超过基线的200%
4. 完成发明专利一项，高水平论文一篇

课题二：硬件加速规则匹配技术

研究背景：随着互联网的快速发展和各行业的数字化转型，网络流量日益增长，对网络安全产品提出了更高的性能要求，以往的基于 CPU 的纯软件实现方案已经很难满足未来的低延迟大流量网络安全需求。于此同时，智能网卡 SmartNIC 处于快速发展期，智能网卡可配备 FPGA、ASIC 等特定芯片，为网络传输和安全检测等负载提供了超高的硬件处理能力。

研究内容：

1. 研究并设计基于智能网卡的规则匹配技术。
2. 研究智能网卡性能优化技术

课题提供：

1. 硬件平台，使用 Intel Xeon CPU 和 BlueField-2 智能网卡
2. 待加速负载，用于测试的规则（300 条规则），用于测试的网络数据流（100Gbps）

考核指标：

1. 使用智能网卡加速的版本和 CPU 纯软件版的匹配结果一致
2. 使用智能网卡加速的版本可以满足线速匹配性能要求，即能够在 100Gbps 的网络流量下完成规则匹配
3. 使用智能网卡加速的版本，规则匹配引入的延迟不超过 20us
4. 完成发明专利一项，高水平论文一篇

课题三：安全威胁样本生成技术

研究背景：本问题一直是 ML 用于安全领域的主要挑战之一，特别是在终端动态入侵检测领域。首先，相较于静态样本的特征数据，动态数据的获取难度极大：目前主流方法通过沙箱或终端探针动态运行获取，前者有沙箱对抗问题，后者有数据粒度不足的问题；其次，样本的多样性难以保证：其一，根据 att&ck 矩阵，攻击技术繁多，杀伤链复杂，难以覆盖全面；其二：目前终端动态入侵检测普遍基于 provenance graph 的数据模型，单进程或执行预制脚本的方式模拟恶意行为，其运行数据在图中结构十分相似，与真实攻击相差甚远；最后，数据分布失衡问题严重：一方面，合法程序的行为数据量巨大，而恶意样本的行为数据缺乏；另一方面，真实环境中攻击态势持续改变、样本分布快速变化，而实验室中构造的训练集样本分布难以契合实际情况。这一些列的问题，导致 ML 在安全领域难以落地。然而，对于未知攻击或 APT 攻击检测，ML 是解决的最佳候选方法之一。因此，如何自动化产生符合实际需求的样本数据具有重要的研究意义。

研究内容：研究和开发 APT 攻击样本生成技术，该技术应可以自动化地构造在执行顺序（杀伤链）、provenance graph 等方面具有多样性的攻击样本，同时要考虑样本的分布、快速迭代等问题，为终端动态行为检测 AI 研究提供基础。

1. 在 Windows/Linux（可选：MacOS）上，研究提供自动化构造样本的方法或者提供产生具备真实性的动态行为数据的方法。样本对应的行为应具备多样性，特别在 provenance graph 的视角上。
2. 研究通过 CTI 等方式，跟踪攻击态势变换，可灵活增加多样性、改变样本分布特征的方法。
3. 孵化并实现相关框架或工具。

课题提供：Windows 和 Linux 上，能够采集系统事件、行为数据（包括进程、文件、注册表、计划任务、网络等）的工具，以及部分恶意脚本。

考核指标：生成具备在 provenance graph 上具备多样性和真实性的攻击样本/框架，生成技术不限于 AI。

1. 提供不少于 30 个不同场景的攻击生成
2. 生成的样本应能覆盖同类已知真实攻击 70%
3. 理论创新方面，能在相关领域发表高价值学术论文。
4. 完成发明专利一项，高水平论文一篇

课题四：基于 WASM 的安全容器研究

课题背景：WebAssembly 是一种新的二进制格式的开放标准。从设计上看，它是内存安全的、可移植的，并以接近原生的性能运行。其他语言的代码可以交叉编译成 WebAssembly。目前，对 Rust、C/C++和 AssemblyScript 提供了一流的支持。许多其他编译器已经在开发中。WASI 是一项新的标准，它将 WebAssembly 的执行扩展到操作系统。它引入了新的抽象层次，使 WASM 二进制文件可以“编译一次，就能在任何地方运行”，而与底层平台无关。WASM 具备两个主要特征：（1）与容器相比，WASM 及其运行时可以快速执行并且体积非常小（2）WASM 在默认情况下不能做任何事情；只有在明确的权限下才能执行。随着时间的推移，WASM 可能会成为最流行的容器类型之一。目前基于 wasm 已有一些开源项目，但是这些项目往往带有很多的实验性质，存在兼容性，性能，安全性等诸多问题，距离实际应用还有不少差距。例如，Inclavare Container 对 X86 芯片具有严格要求，兼容性较差；RLBox 代码迁移需要耗费较高人力，实施性较弱。因此，本课题希望建立一整套较为完备的基于 wasm 技术的安全容器环境构建方案。

研究内容：

1. 分析 wasm 容器对不同芯片的兼容模式与问题，在保证对 x86 下主流 Linux 的兼容性的前提下，对 wasm 容器迁移到其他平台的功能和性能进行测试
2. 研究和分析不同开源项目在组织形式、功能类型和性能指标等方面的优劣性，并设计性能优化方案。
3. 研究 wasm 容器的安全性测试方案，对 wasm 容器进行全面的安全分析，包括 wasm 容器的安全分析以及相关测试方法。

课题提供：服务器资源，以及安全测试工具

考核指标：

1. 提供兼容性，性能和安全测试方案及结果
2. 基于测试结果，提供一种性能提升方案
3. 提供满足性能要求的安全容器设计方案，方案内容包括但不限于容器运行时、镜像打包、容器部署等
4. 完成发明专利一项，高水平论文一篇