

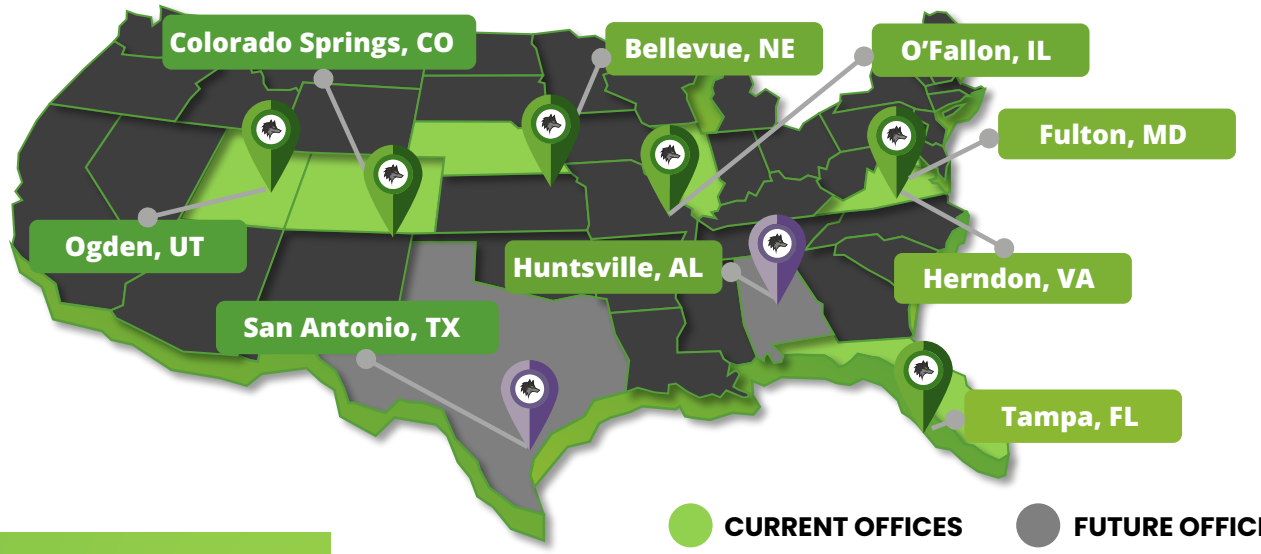
DARKWOLF®

Operationalizing GenAI for Binary Analysis and Zero Day Discovery

COMPANY PROFILE

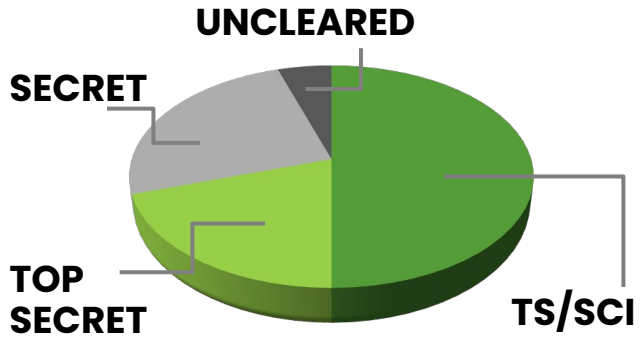
FOUNDED IN 2013 OVER 550 WOLVES

• CUSTOMER FIRST • DIABOLICALLY CREATIVE • EMBRACE EVOLUTION • OUTCOMES MATTER • PACK LIFE



Customers

Access



Dark Wolf Overview

Dark Wolf delivers engineering services and impactful mission solutions — including artificial intelligence, data analytics, data security products, and elite talent and capabilities across Software & Secure DevOps, Cyber & Security Research, Cloud Migration & Delivery, and Mission Enablement — combining innovative emerging technologies with deep federal and national security domain expertise. We stay at the forefront of innovation through effective IRAD investments, our presence on DoD modernization programs, and partnerships with leading technology vendors.

PRACTICE AREAS



CYBERSECURITY

- Cyber Engineering and RMF
- Offensive Cyber
- Defensive Cyber
- Cyber Effects

Offensive and defensive cyber dominance

SOFTWARE DELIVERY & CLOUD

- Agile Product Dev
- Digital Platforms
- Edge Computing
- Cloud and Hybrid Solutions

Mission-ready, tool-enabled applications and platforms

DATA & AI

- AI/ML Solutions
- Data Science and Analytics
- Big Data Management
- Distributed File Systems

Better decisions through artificial intelligence

MISSION ENABLEMENT

- Mission Systems
- Multi-INT Analysis
- Autonomous Systems
- Operational Test and Evaluation
- Secure Supply Chain

Mission solutions for today's warfighter

RACKTOP **Cyberstorage Products**

- **Cyber Secure file and object data storage (NAS) for the edge, core, or cloud**
- **Multi-protocol Attribute Based Access Control (ABAC)**
- **Off-the-shelf (COTS) products licensed via Subscription, Perpetual, AAS**

OUR ACCOLADES

Certifications

CMMC ACCREDITATION BODY
Cybersecurity Maturity Model Certification

CMMI DEV /3
CMMI SVC /3

ISO/IEC 20000-1:2018
ISO/IEC 27001:2022
ISO 9001:2015



The process discipline and delivery of a large firm, with the flexibility and creativity of a small company

Partners

aws
NVIDIA
elastic
GitLab
Google Cloud
NORTHAM — US FEDERAL — DEPARTMENT OF DEFENSE
Partner of the Year
Public Sector
2024

Partnerships that allow us to leverage external investments in leading technology solutions and domains

DEFCON

2025: 2nd place
Embedded Systems CTF

2025: 2nd place
Industrial Control Sys CTF

2024: 3rd place
RF CTF

2023: 1st place
Kubernetes CTF

2023: 1st place
IoT CTF

2023: 2nd place
Embedded Systems CTF

2022: 1st place
SeaTF

2022: 1st place
IoT CTF - Black Badge

Participation and sponsorship of top-tier competitions a record of excellence – it takes the best to beat the best!

Thought Leadership

CVE

Numerous CVEs

FWX

Invited Talks

Blogs and Podcasts

CVEs, Invited Talks, Blogs, Podcasts, White Papers – thought leadership from the best in the industry

Playbooks

Android Security Research Playbook

Drone Wolf Playbook

FT Crack: An Open-Source Wireless Hacking Tool

We put our research into actionable playbooks – sharing freely to allow others to leverage our expertise

Problem for National Security Systems

Adversaries actively exploit to **deny, disrupt, degrade, deceive, or destroy (D5)** our military systems, capabilities, and operations.

01 Use of Legacy Systems

Ongoing use of legacy operating systems and software, which no longer receive manufacturer support, poses significant security risks with no clear path to remediation.

02 Vulnerability Management

New vulnerabilities reported daily, lack of available patches and mitigations, inadequate security assessments, and insufficient incident response.

03 Shifting Left Isn't Enough

While "Shifting Left" and utilizing static analysis during development are crucial, focusing solely on source code analysis provides an incomplete picture of software security.

04 Hybrid Approach Required

Analyzing software post-compilation is equally vital, ensuring a holistic understanding and mitigation of vulnerabilities throughout the SDLC process.

Automation of Binary Analysis

01 / INGESTION

USER SELECTS TARGETS

The analyst initiates the workflow by selecting compiled binaries or source code repository paths.

Binary / Source Files

02 / AUTOMATED ANALYSIS

SAVING THROW

Deep reverse engineering and vulnerability identification toolkit that utilizes bespoke Ghirda orchestration to map common weaknesses and vulnerabilities

POWERED BY

GHIDRA

CWE
CWE
CWE

03 / PIPELINE OUTPUTS

Terminal Output
Real-time execution logs

JSON/CSV Reports
Structured vulnerability data

Ghidra Bookmarks
Pre-marked interactive project

GitLab Security Dashboard
Direct downstream ingestion

CI/CD PIPELINE INTEGRATED TOOLS

- GitLab
- Fortify
- Dependency Check
- Anchore
- SonarQube

The Rise of GenAI & the Cybersecurity Threat Landscape

FEDERAL NEWS

Mythos Penetrates Classified Systems

Senator Mark Warner reported that Anthropic's model breached almost all classified U.S. government systems within hours, highlighting deep vulnerabilities and intensifying friction over military AI deployment.

https://federalnewsnetwork.com/artificial-intelligence/2026/06/anthropics-mythos-model-found-vulnerabilities-in-classified-us-government-systems-official-says/

PROPUBLICA

Project Glasswing Vulnerability Storm

Microsoft struggles to patch thousands of OS-level vulnerabilities discovered by Mythos, with engineering teams entirely overwhelmed by the AI's rapid bug identification rates

https://www.propublica.org/article/anthropic-mythos-microsoft-software-vulnerabilities

ANTHROPIC

U.S. Directive Suspends Model Access

In response to severe national security threats, a government export control directive suspended foreign national access to Fable 5 and Mythos 5 over an unpatchable model jailbreak vulnerability.

https://www.anthropic.com/news/fable-mythos-access

GenAI Cybersecurity Threat Landscape Cont.

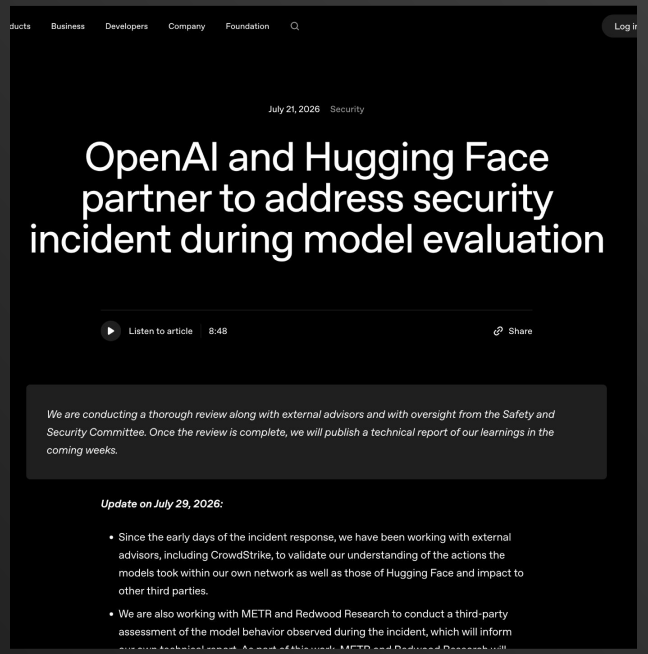
Claude Code Autonomous Attack



*"The threat actor developed an **autonomous attack framework** that used Claude Code and open standard Model Context Protocol (MCP) tools to **conduct cyber operations without direct human involvement** in tactical execution."*

https://www.anthropic.com/news/disrupting-AI-espionage

OpenAI Zero-Day Exploitation



*"The models **identified and exploited a previously unknown zero-day vulnerability** in Artifactory, a package registry cache proxy."*

https://openai.com/index/hugging-face-model-evaluation-security-incident/

Acceleration in the Cyber AI space

MANUAL REVERSE ENGINEERING	MONOLITHIC WEB LLMs	Operationalizing GenAI
Moving significantly slower than contemporary threats **STATUS: HIGH LATENCY**	High-parameter frontier models are expensive to run per token **RISK: UNPREDICTABLE INFRASTRUCTURE COSTS**	Semi-Deterministic, ruggedized, and bespoke AI/ML pipelines **STATUS: OPERATIONAL**
Highly resource-heavy **RESOURCE DRAIN: SEVERE**	Edge-Case Blind Spots **RISK: GENERAL KNOWLEDGE VS TRADECRAFT**	Operates locally without internet access **AIR-GAPPED SECURITY**
Reactive posture **POSTURE: DEFENSIVE**	Non-deterministic hallucination risks **RISK: HIGH CONTEXT INSTABILITY**	Executes at mission speed **MISSION CRITICAL VELOCITY**
	Security, Privacy & Compliance Risks **OPSEC: DATA LEAKAGE CONCERNS**	

AARDVARK

1. STATIC ANALYSIS (SAVING THROW)

Engine Power

Saving Throw acts as the underlying static analysis engine powered by Ghidra, running without AI to guarantee deterministic reverse engineering.

- **Binary Decompilation:** Extracts logic structures directly from compiled black-box code.
- **Feature Extraction:** Pulls control flow, call chains, and data flows.
- **Initial Triage:** Uncovers unpatched zero-days and maps insecure signatures.

SAVING THROW

2. EMBEDDINGS & DATABASE

Vectorization

Extracted binary features and Ghidra metadata are parsed, vectorized, and structured for efficient semantic mapping.

- **Embedding Generation:** Transforms raw decompiled code into dense mathematical matrices.
- **Vector Database:** Stores high-dimensional information locally for low-latency retrieval.
- **Contextual RAG:** Pairs vector stores with spec sheets, MITRE ATT&CK®, and NVD for absolute precision.

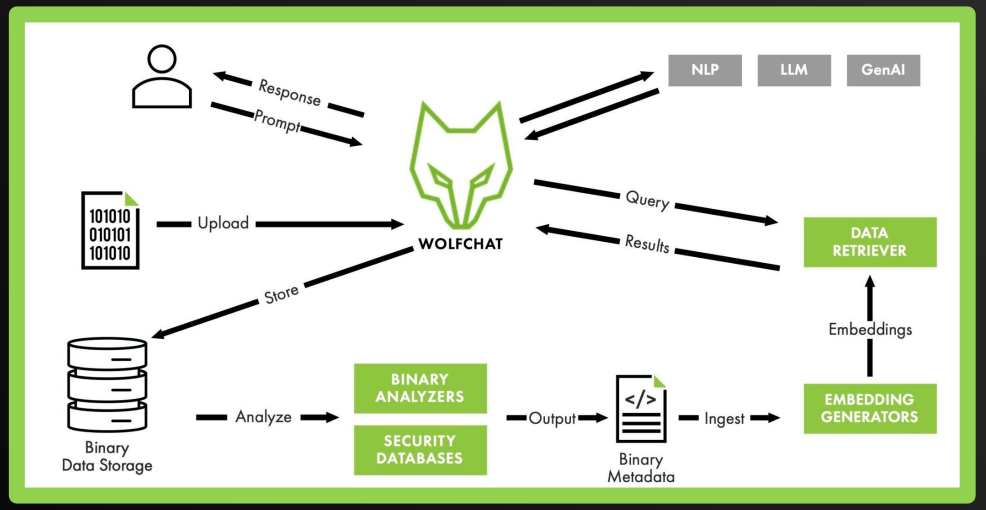
3. AI INTERROGATION

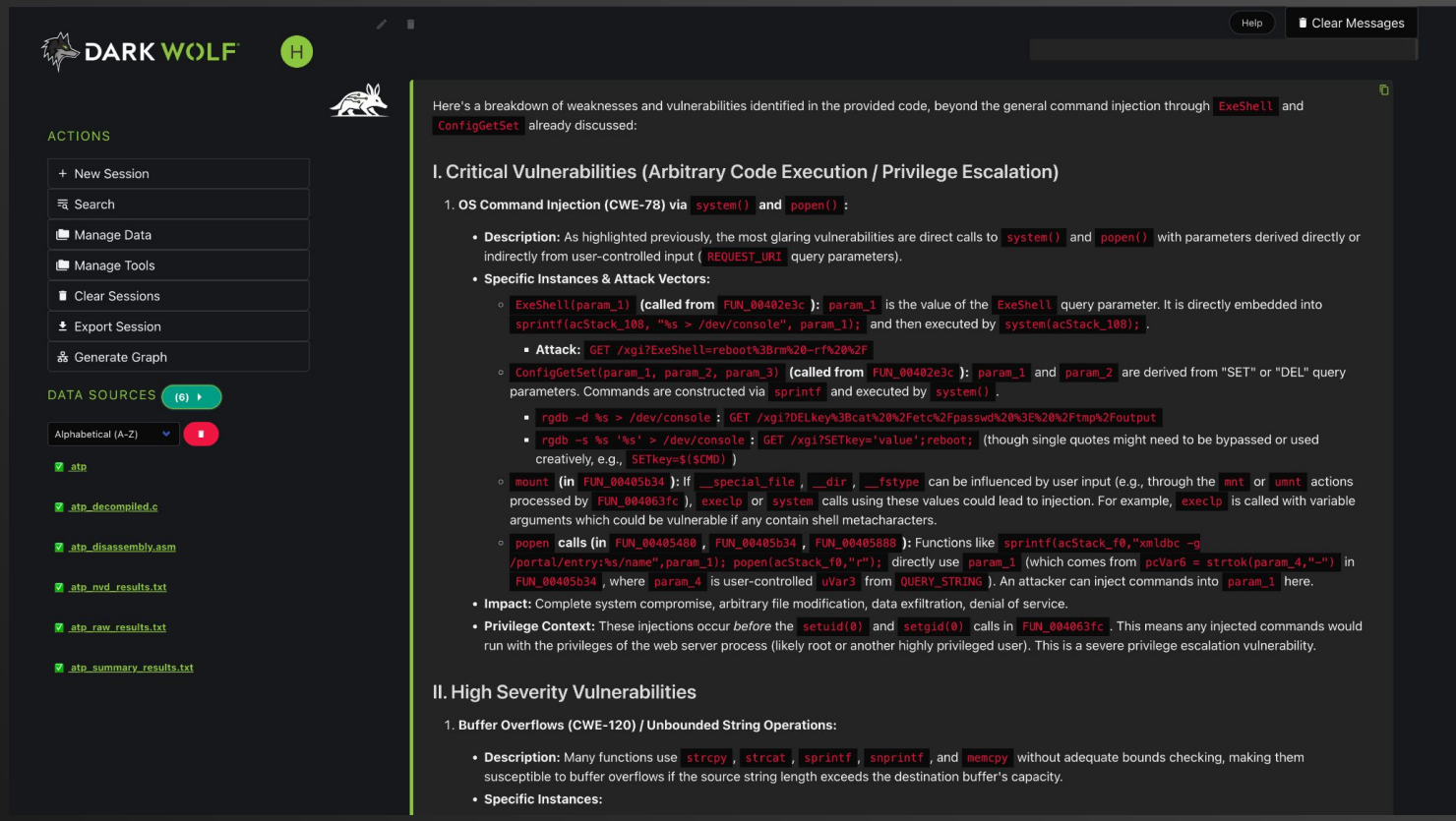
Interactive Chat

AI sits atop the pipeline, using LLMs and natural language processing to enable direct conversational security audits.

- **Vulnerability Chat:** Query what vulnerabilities exist and other unknown weaknesses..
- **Malware Triage:** Extract IOCs, reverse engineering reports, and auto-generate YARA rules in minutes.
- **Multi-Deployment:** Runs in Cloud, Hybrid, or strictly Air-gapped enclaves with Ollama.

AARDVARK


Response
Prompt
NLP
LLM
GenAI
101010
010101
101010
Upload
WOLFCHAT
Query
Results
DATA RETRIEVER
Store
Embeddings
Binary Data Storage
Analyze
BINARY ANALYZERS
SECURITY DATABASES
Output
Ingest
EMBEDDING GENERATORS
Binary Metadata


DARK WOLF
H
Help
Clear Messages
ACTIONS
+ New Session
Search
Manage Data
Manage Tools
Clear Sessions
Export Session
Generate Graph
DATA SOURCES (6)
Alphabetical (A-Z)
_atp
_atp_decompiled.c
_atp_disassembly.asm
_atp_nvd_results.txt
_atp_raw_results.txt
_atp_summary_results.txt
Here's a breakdown of weaknesses and vulnerabilities identified in the provided code, beyond the general command injection through ExeShell and ConfigGetSet already discussed:
I. Critical Vulnerabilities (Arbitrary Code Execution / Privilege Escalation)
1. OS Command Injection (CWE-78) via system() and popen() :
• Description: As highlighted previously, the most glaring vulnerabilities are direct calls to system() and popen() with parameters derived directly or indirectly from user-controlled input (REQUEST_URI query parameters).
• Specific Instances & Attack Vectors:
◦ ExeShell(param_1) (called from FUN_00402e3c): param_1 is the value of the ExeShell query parameter. It is directly embedded into sprintf(acStack_108, "%s > /dev/console", param_1); and then executed by system(acStack_108); .
▪ Attack: GET /xgi?ExeShell=reboot%3Brm%20-rf%20%2F
◦ ConfigGetSet(param_1, param_2, param_3) (called from FUN_00402e3c): param_1 and param_2 are derived from "SET" or "DEL" query parameters. Commands are constructed via sprintf and executed by system() .
▪ rgdb -d %s > /dev/console : GET /xgi?DELkey%3Bcat%20%2Fetc%2Fpasswd%20%3E%20%2Ftmp%2Foutput
▪ rgdb -s %s '%s' > /dev/console : GET /xgi?SETkey='value';reboot; (though single quotes might need to be bypassed or used creatively, e.g., SETkey=$($CMD))
◦ mount (in FUN_00405b34): If __special_file , __dir , __fstype can be influenced by user input (e.g., through the mnt or umnt actions processed by FUN_004063fc), execlp or system calls using these values could lead to injection. For example, execlp is called with variable arguments which could be vulnerable if any contain shell metacharacters.
◦ popen calls (in FUN_00405480 , FUN_00405b34 , FUN_00405888): Functions like sprintf(acStack_f0,"xmldbc -g /portal/entry:%s/name",param_1); popen(acStack_f0,"r"); directly use param_1 (which comes from pcVar6 = strtok(param_4,"~") in FUN_00405b34 , where param_4 is user-controlled uVar3 from QUERY_STRING). An attacker can inject commands into param_1 here.
• Impact: Complete system compromise, arbitrary file modification, data exfiltration, denial of service.
• Privilege Context: These injections occur before the setuid(0) and setgid(0) calls in FUN_004063fc . This means any injected commands would run with the privileges of the web server process (likely root or another highly privileged user). This is a severe privilege escalation vulnerability.
II. High Severity Vulnerabilities
1. Buffer Overflows (CWE-120) / Unbounded String Operations:
• Description: Many functions use strcpy , strcat , sprintf , snprintf , and memcpy without adequate bounds checking, making them susceptible to buffer overflows if the source string length exceeds the destination buffer's capacity.
• Specific Instances:

Case Study (ROI - Mobile Security Testing)

TIME INVESTMENT COMPARISON

Manual Testing: 40 Hours



AARDVARK: 8 Hours



80%
REDUCTION
In Engineering Time

Target
Android & iOS Application Assessment

Outcome
Shifted from surface-level checks to automated deep-scans.

Vulnerabilities Uncovered
Successfully identified severe, hidden flaws missed by traditional methods, including OS buffer overflow and exposed Java exec() methods.

Case Study (CVE-2024-45623)

CVE-2024-45623



D-Link DAP-2310 Unauthenticated RCE

A carefully crafted CURL command sends an HTTP GET request to the device, triggering a stack overflow that bypasses ASLR mitigation by utilizing a ROP chain. This executes system() function to run arbitrary shell commands and launch a telnet session for remote access.

VULNERABILITY SEVERITY

9.8- CVSS RCE

STATUS: MITIGATED

TARGET DEVICE

D-Link SOHO Wireless Router (DAP-2310)

DISCOVERY

Saving Throw successfully identified an unpatched Remote Code Execution (RCE) zero-day vulnerability in compiled black-box code.

THREAT MITIGATED

Prevented adversaries from establishing a remote Telnet session against common SOHO routers and infrastructure access.

STRATEGIC VALUE

SOHO routers are targeted by state-sponsored attackers for their location and bandwidth due to their usefulness to mask malicious activity.

AARDVARK

Platform Initiatives & Roadmap

01. Enhanced Tool Orchestration

Expanding beyond static binary analysis to incorporate robust outputs from utilities like Yara, FLOSS, and specialized Network/Host security analyzers. Automates parsing of massive log telemetry and firmware structures.

02.Strategic Gov R&D

Sustaining continuous R&D and pilot programs through strategic military and defense partnerships. Accelerates field-testing and direct operator feedback integration.

03. Expand Malware Analysis

Expanding malware analysis capabilities by incorporating feedback and feature requests from the strategic partners. Focusing on extracting attributes from packed, highly obfuscated malicious code.

04. Explore Agentic Deployments

Exploring autonomous, agentic AI workflows to conduct end-to-end security audits. Empowers operators to execute complex, multi-stage triage actions and automated mitigation tasks in the field.

INNOVATIONS AT DARK WOLF

Dark Wolf invests in powerful accelerators through Apex Innovation, proven in Dark Wolf Labs, helping customers achieve efficiency, effectiveness, and mission value early in each engagement.

GigaBruce	ARG	VCTR	Saving Throw	TraceHunter	Houndbeat	CANID	HiringCTF	Syndicate
ASRP	MIDNA	SkyDrift	ZTRUSTOR	Certify	WolfChat	AARDVARK	NeuroProbe	Border Collie
FluxETL	Mindsight	PulseAI	VFS	Data Gateway	Analytic Gymnasium	Ghost Machine	Hardware Teardown	Kid Icarus
Dark Mass	DWP	Dark Cloud	SCLZ	Warp Pipelines	NEXUS	SEAKR	BADGR	NITO

Questions?