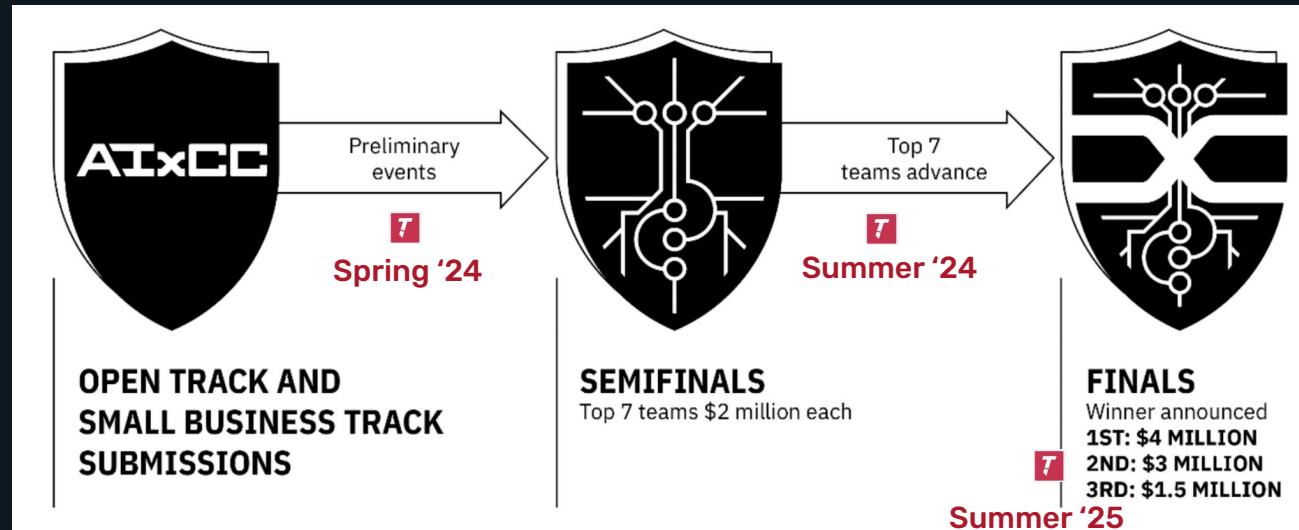


Buttercup and DARPA's AI Cyber Challenge

AI Cyber Challenge (AIxCC)

Competition to create AI systems that find and fix bugs in open-source software



Competition Rules



- **Fully automated solution: no human in the loop**
- **Points for:**
 - Patches - Worth the most points
 - Vulnerabilities - Requires input which triggers bug
 - Static analysis alerts - Minor points
 - Bundling - Match patches, vulnerabilities, and alerts
- **Scoring modifiers:**
 - Speed - Earlier submissions get more points
 - Accuracy - Incorrect/duplicate submissions reduce points
- **Budget**
 - \$50,000 LLM API
 - \$85,000 Azure

How would you solve this?

Results: \$3 Million Prize

Buttercup won 2nd place and a \$3 Million prize

Keys to success:

- Accuracy
- Versatility
- Reliability



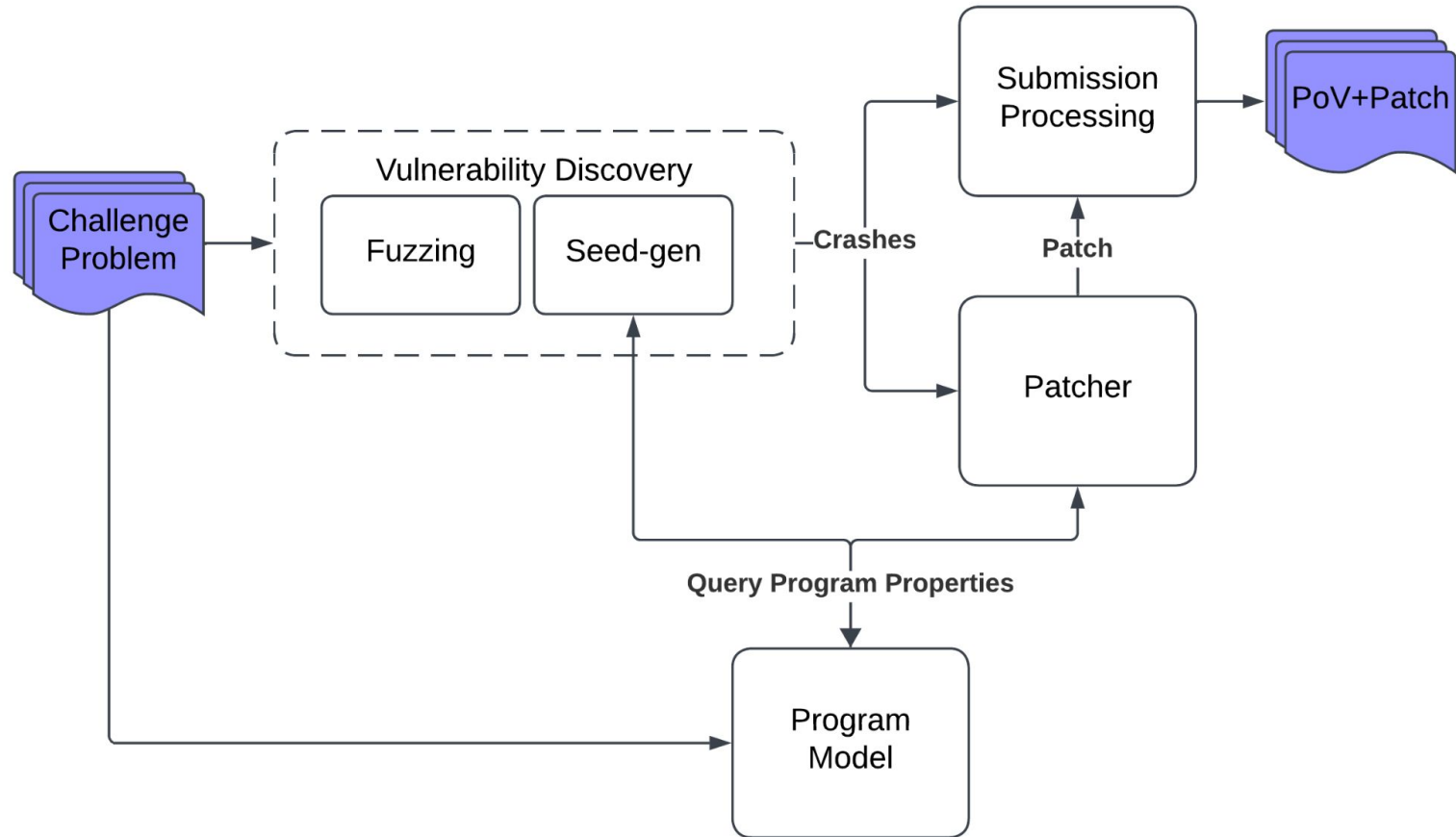


Team	LLM spend	Compute spend	Total spend	Cost per point
Team Atlanta	\$29.4k	\$73.9k	\$103.3k	\$263
Trail of Bits	\$21.1k	\$18.5k	\$39.6k	\$181
Theori	\$11.5k	\$20.3k	\$31.8k	\$151
fuzzing_brain	\$12.2k	\$63.2k	\$75.4k	\$490
Shellphish	\$2.9k	\$54.9k	\$57.8k	\$425
42-b3yond-6ug	\$1.1k	\$38.7k	\$39.8k	\$379
LACROSSE	\$631	\$7.1k	\$7.2k	\$751

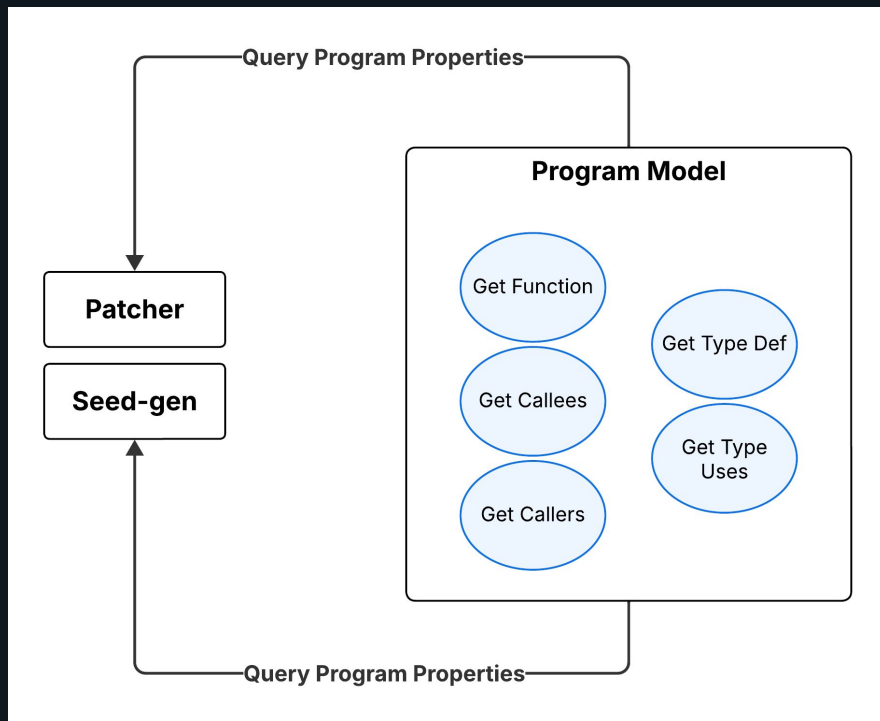
Results!

2nd in LLM Spend
6th in Compute Spend
2nd in \$ per Point

Buttercup System Design



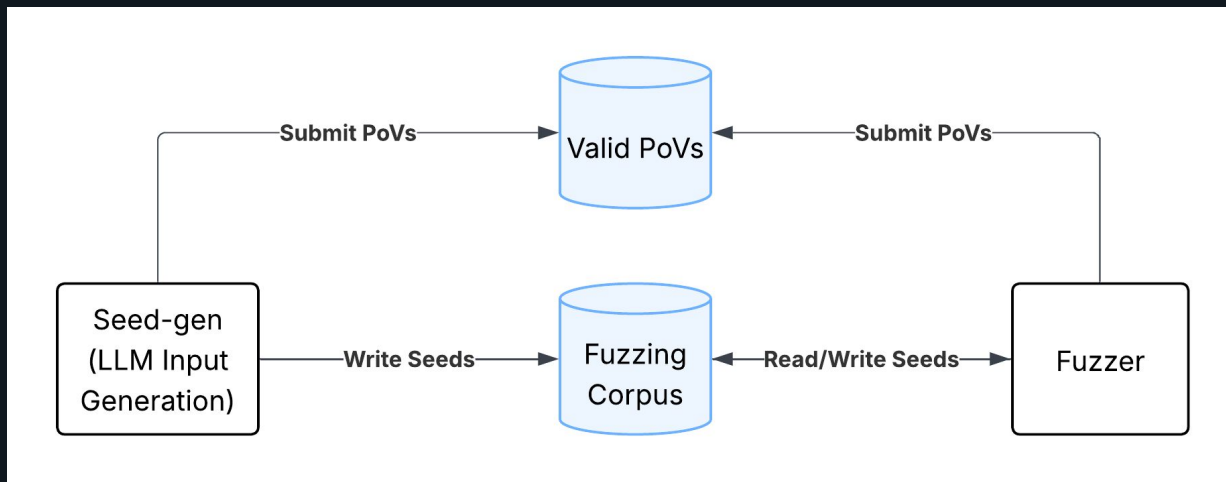
Program Model



- Constructs program model using CodeQuery + Tree-sitter
- Supports querying program properties (functions & types)
- Used by LLM components

Vulnerability Discovery

Approach: Combine fuzzing and LLM input generation



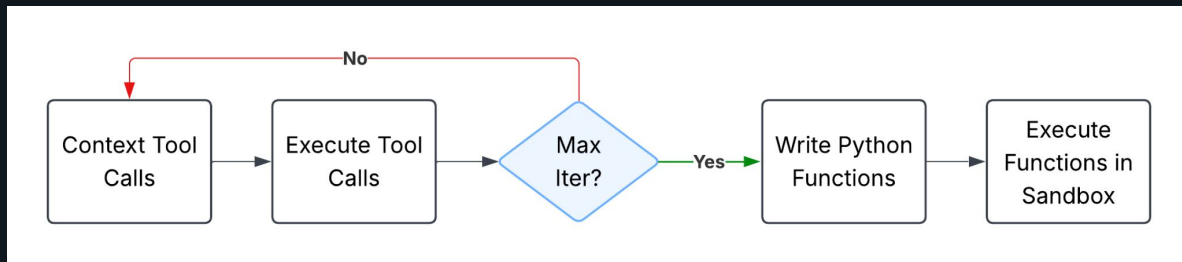
Note: PoVs stands for Proofs of Vulnerability

Fuzzing

- **Standard OSS-Fuzz fuzzers:**
 - LibFuzzer for C/C++
 - Jazzer for Java
- **Fuzzer bots sample active harnesses to run short fuzz campaigns**
- **Merger bots save inputs which improve coverage**
 - Merge a fuzzer bot's local corpus to the shared corpus

Seed-Gen: Enhancing Fuzzing

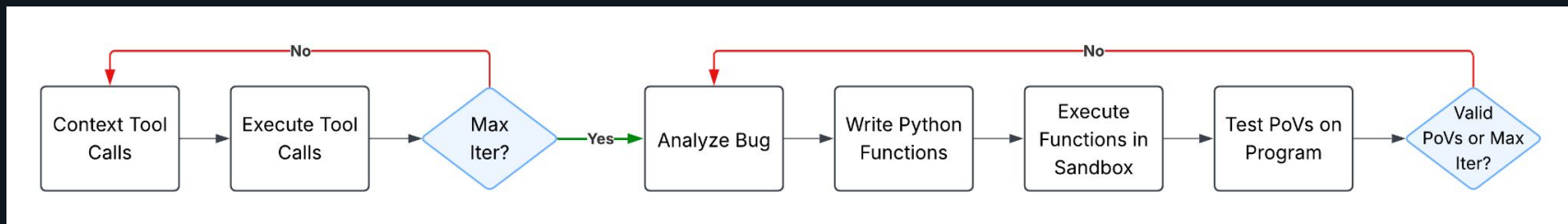
- Generate inputs with LLM to support the fuzzer
- **Initialize Task:** Bootstrap fuzzer with initial seed inputs that exercise harness
- **Explore Task:** Increase coverage for a target function
 - Sample function with low coverage
 - Generate reaching inputs



Initialize and Explore Tasks

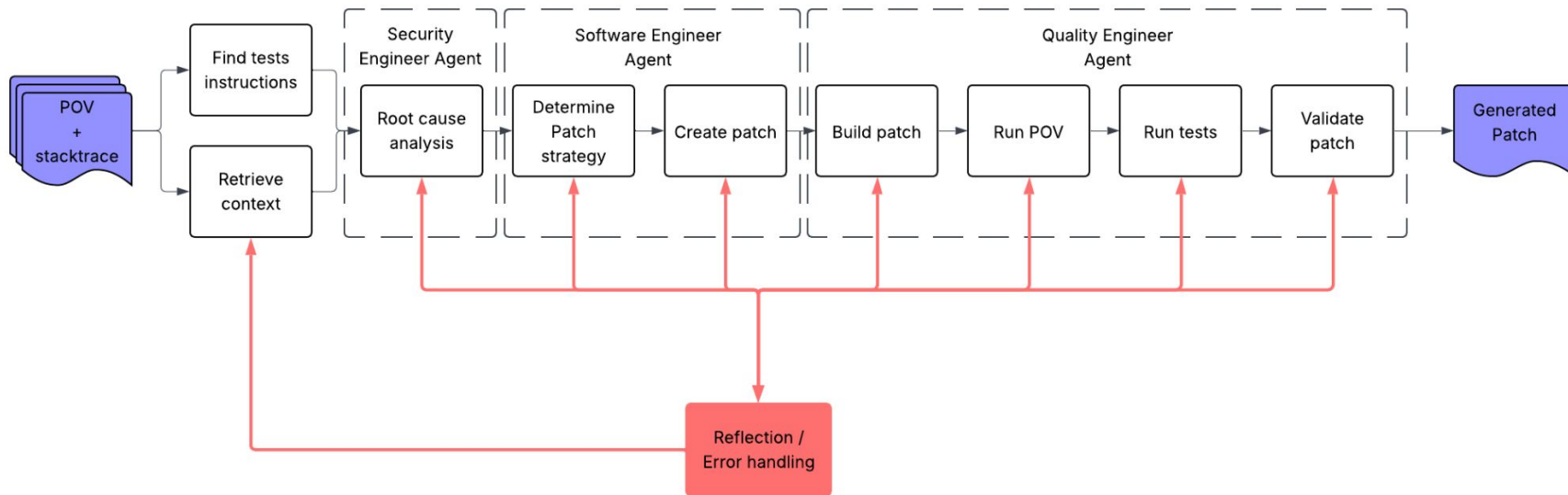
Seed-Gen: Vulnerability Discovery

- **Independently find bugs with an LLM agent**
- **Vuln discovery task: Identify vulnerabilities and create PoVs**
 - Similar to fuzzing

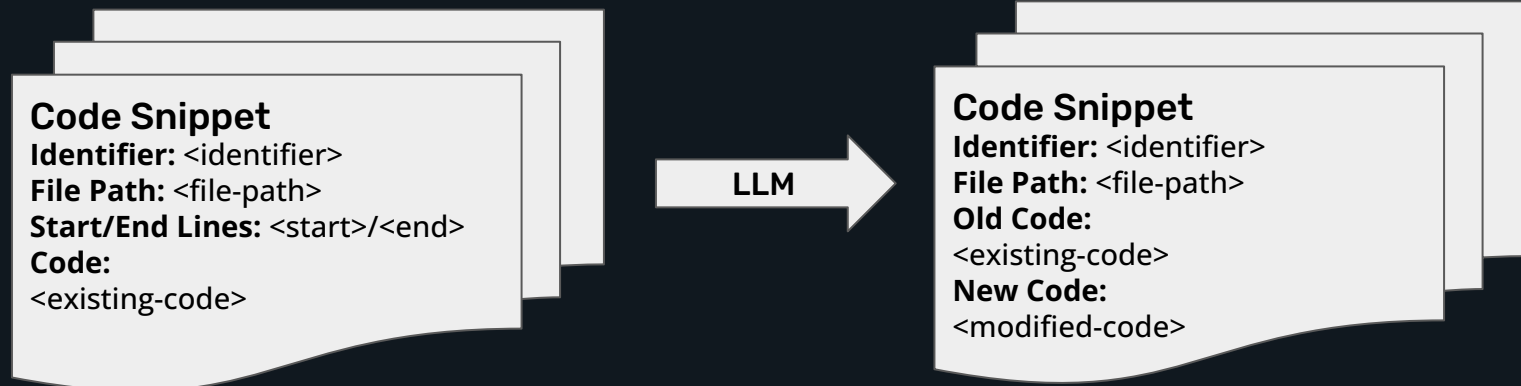


Vulnerability Discovery Task

Patcher Design



Patch Creation



Using Buttercup

Repo: github.com/trailofbits/buttercup

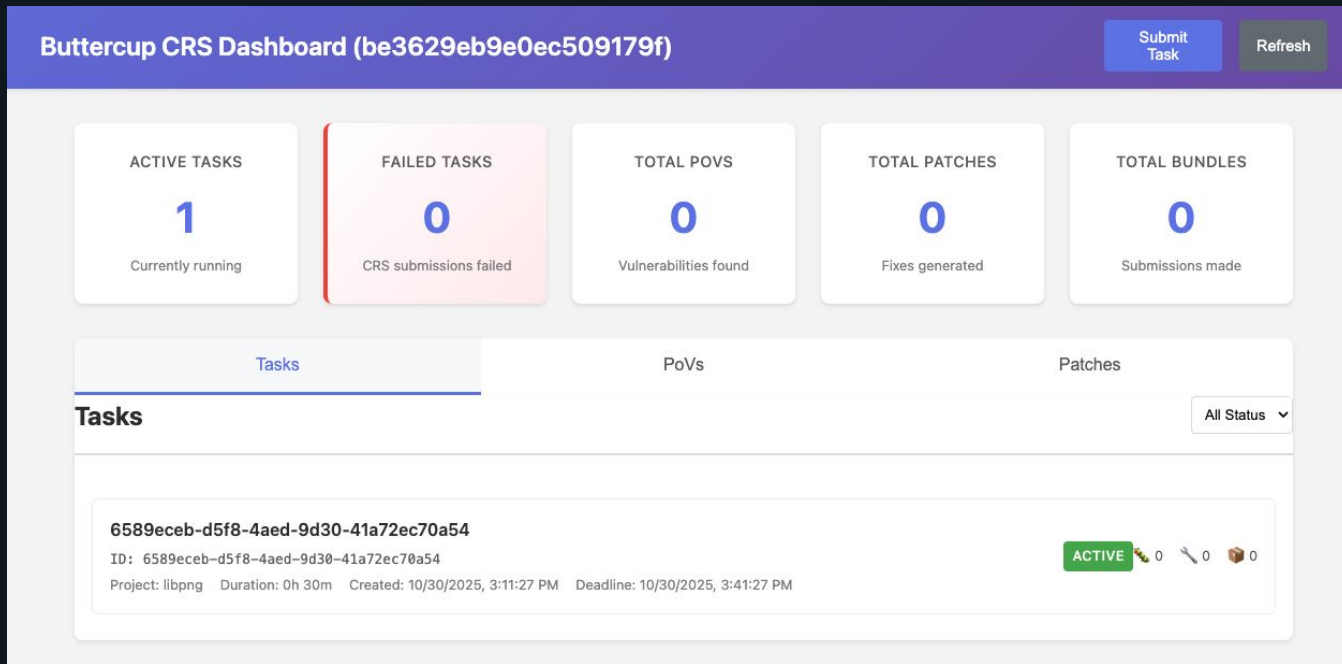
Requires:

- LLM API Key (OpenAI, Anthropic, or Gemini)
- OSS-Fuzz compatible project as a target

Easy to use:

- Seamless configuration script
- Web UI to task Buttercup and monitor results

Web UI



Web UI - PoV Review

Buttercup CRS Dashboard (fb40ddc2b222cb90a148) Submit Task Refresh

POV: 749b444a-3465-4388-bfe4-6b6d9abcc81b

Artifact Information

ID: 749b444a-3465-4388-bfe4-6b6d9abcc81b

Task: libpng

Task ID: 446609ae-e6e8-4f98-af60-4523471dbe15

Timestamp: 11/3/2025, 10:31:24 AM

Architecture: aarch64

Engine: libfuzzer

Fuzzer: libpng_read_fuzzer

Sanitizer: address

Testcase Size: 173 bytes

Testcase Preview:

```
00000000 89 50 4e 47 0d 0a 1a 0a 00 00 00 0d 49 48 44 52 [PNG.....IHDR]
00000010 00 00 00 02 00 00 00 01 01 03 00 00 00 ce 49 48 [.....IHDR]
00000020 44 52 00 00 04 69 43 43 50 67 77 b6 b7 73 42 49 [IDR...CCPgw...aBI]
00000030 54 97 57 28 52 52 52 b6 e6 df de de de 4e 02 68 [tW(RRR,e...N,h]
00000040 69 68 50 e9 54 a2 c7 8f d5 6a 75 f7 e6 cd 5a 04 [bMP....3a...].]
00000050 a1 a0 e9 8d d7 af 47 7b bd 37 d3 d3 95 69 69 41 [..M..G|7...iSA]
00000060 4d 41 00 01 00 00 01 45 7e 00 65 58 49 66 63 48 [MA....E..eXfch]
00000070 52 4d 00 00 7a 26 00 00 01 00 00 3a 98 00 69 43 [RM..s6.....IC]
... (45 more bytes)
```

Download POV

Error: failed to submit task to CRS (HTTP status: 500)

Web UI - Patch Approval

PATCH: a60444b9-607f-477c-8e89-6cae445e2cf0

Artifact Information

ID:

a60444b9-607f-477c-8e89-6cae445e2cf0

Task:

example-libpng

Task ID:

2ce9eafc-bb91-4583-9e23-bfb0fdd41adc

Timestamp:

11/3/2025, 5:11:19 PM

Status:

ACCEPTED

Patch Size:

2536 characters (1902 decoded)

Patch Content:

```
diff --git a/pngutil.c b/pngutil.c
index 01e08bf..76b3766 100644
--- a/pngutil.c
+++ b/pngutil.c
@@ -1443,20 +1443,21 @@ png_handle_iCCP(png_structp png_ptr,
png_inforp info_ptr, png_uint_32 length)
{
    keyword_length = 0;
-   while (keyword_length < (read_length-1) && keyword_length <
read_length &&
+   while (keyword_length < max_keyword_wbytes-1 &&
```

Download PATCH

Approve Patch

Reject Patch



Buttercup

**TRAIL
OF
BITS**

Try out Buttercup:

github.com/trailofbits/buttercup

Website: trailofbits.com

Careers: trailofbits.com/careers

Blog: blog.trailofbits.com