End-to-End Verification of Stack-Space Bounds for C Programs

Quentin Carbonneauax
Tahina Ramananandro
Jan Hoffmann
Zhong Shao

Yale University

April 14th, 2014
Does this program safely run?

```c
#include <stdint.h>
typedef uint64_t t;

void f (t* pa, t* pb) {
    if (*pa == 0) return;
    *pa--;
    f (pa, pb);
    *pb++;
}

int main (int argc, char* argv[]) {
    t a = UINT64_MAX, b = 0;
    f (&a, &b);
    return a;
}
```

- gcc -O0 && ./a.out
  - Segfault (stack overflow)
- gcc -O1 && ./a.out
  - OK (function inlining)
Toyota's killer firmware: Bad design and its consequences

Published on: October 29, 2013

By Mark Fuge

Toyota's Toyota's report for 2013 of its efforts to improve the design of its electronic control units (ECUs) highlighted a number of critical design flaws that led to the death of one of its employees. The report, which was released last week, focused on the ECU Central Failure Analysis Report (CFAR).

The CFAR highlighted various issues with the design of the ECU that contributed to the tragic accident. The report stated that the design was not robust enough to handle unexpected situations and that the ECU's software was not well-integrated with the vehicle's overall system.

Toyota's design flaw case study is a cautionary tale for other automotive manufacturers to ensure that their designs are thoroughly tested and validated to prevent similar incidents in the future.

The report also called for the establishment of a new committee to oversee the design and testing of ECUs and other automotive systems. The committee would be charged with ensuring that designs are safe and reliable and that they meet the highest standards of safety and quality.

Toyota's CFAR has been widely criticized for its lack of transparency and for not adequately addressing the root causes of the issue. The company has faced intense scrutiny for its handling of the incident, and questions have been raised about its efforts to prevent similar incidents in the future.

Toyota's efforts to improve the design of its ECUs have been met with mixed reviews. While the company has made some progress, there is still much work to be done to ensure that its designs are safe and reliable.

To prevent similar incidents from occurring in the future, automotive manufacturers must take a proactive approach to design. This includes investing in robust testing and validation processes, collaborating with experts in the field, and establishing strong safety standards.

Toyota's CFAR case study serves as a powerful reminder of the importance of design and testing in the automotive industry. The company must take the necessary steps to address the issues highlighted in the report and work towards creating a safer future for all drivers.
Toyota's killer firmware: Bad design and its consequences

Michael Dunn - October 28, 2013
109 Comments

On Thursday October 24, 2013, an Oklahoma court ruled against Toyota in a case of unintended acceleration that lead to the death of one of the occupants. Central to the trial was the Engine Control Module's (ECM) firmware.

Stack overflow. Toyota claimed only 41% of the allocated stack space was being used. Barr's investigation showed that 94% was closer to the truth. On top of that, stack-killing, MISRA-C rule-violating recursion was found in the code, and the CPU doesn't incorporate memory protection to guard against stack overflow.

Although Toyota had performed a stack analysis, Barr concluded the automaker had completely botched it. Toyota missed some of the calls made via pointer, missed stack usage by library and assembly functions (about 350 in total), and missed RTOS use during task switching. They also failed to perform run-time stack monitoring.
Does this program stack-overflow?

• Important in embedded software
  – led to deadly software bugs in Toyota cars
• Most stack analysis tools available for compiled code only
  – Harder to analyze
  – User interaction is troublesome
• How to prove, at the source level, that the compiled code does not stack-overflow?
  – How to model stack overflow at the source level?
  – How to prove stack-aware compiler correctness?
CompCert

• Formal C and assembly semantics
• Verified semantics-preserving compiler
  – Safety is preserved
  – For safe programs, I/O events and termination/divergence are preserved
[...] it is **hopeless** to prove a stack memory bound on the source program and expect this resource certification to carry out to compiled code: stack consumption, like execution time, is a program property that is not preserved by compilation.
[..] it is **hopeless** to prove a stack memory bound on the source program and expect this resource certification to carry out to compiled code: stack consumption, like execution time, is a program property that is not preserved by compilation.

Really?

Xavier Leroy  
(1968- )
Our solution: Quantitative CompCert

- Introduce stack consumption in C semantics
- Preserve stack consumption by compilation passes: quantitative refinement
- Refine assembly semantics with finite stack
- Make compiler correctness depend on source-level stack bound
  - Introduce a program logic on Clight to derive stack consumption bound
  - Introduce automatic stack analyzer to automatically use program logic on programs without recursion
Overview

Source Program $s$

Verified Quantitative Hoare Logic

Weight Bound $\beta : (\mathcal{E} \rightarrow \mathbb{Z}) \rightarrow \mathbb{N}$

$\forall M : \mathcal{W}_M(s) \leq \beta(M)$

Target Program $C(s)$

Event Metric $M_s : \mathcal{E} \rightarrow \mathbb{Z}$

Stack-Usage Bound $\text{stack}(C(s)) \leq \beta(M_s)$

Safety Proof
Stack consumption in C semantics

- CompCert C produces an I/O event trace
  - Preserved by compilation
- Add function call/return events

- Model the stack consumption as trace weight parameterized by an event metric for call/return events
  - Preserve the weights
  - Stack consumption of a function is parameterized by the stack frame sizes of its callees

- Operational semantics does not go wrong on stack overflow
  - Does not know the event metric, only generates events
Example

```c
int f (int x) {
    return x+1;
}

main () {
    f(0);
}
```

- main() generates trace:
  - call(main) ::
  - call(f) :: return(f) ::
  - return(main) :: nil

- Stack consumption:
  - \(M(main) + M(f)\)

where \(M\) is an event metric (giving non-negative stack frame size for each function)
Stack consumption

- Events $e ::= \ldots \mid \text{call}(f) \mid \text{return}(f)$
- Event and trace valuation:
  \[ V_M(\text{call}(f)) = M(f); \quad V_M(\text{return}(f)) = -M(f); \]
  \[ V_M(e) = 0 \text{ otherwise} \]
  \[ V_M(\text{nil}) = 0; \quad V_M(e::t) = V_M(e) + V_M(t) \]
- Trace weight:
  \[ W_M(T) = \sup \{ V_M(t) \mid T = t \cdot T' \} \]
Stack consumption

Coq implementation: I/O events have constant (maybe non-null) stack consumption

- Event and trace valuation:
  \[ V'_M(e) = V_M(e) \text{ for call/return} \]
  \[ V'_M(nil) = 0; \quad V'_M(t++e::nil) = \max(V'_M(t), V_M(t)+V'_M(e)) \]

- Trace weight:
  \[ W'_M(T) = \sup \{ V'_M(t) | T = t . T' \} \]
Quantitative refinement

For any target behavior $T'$, there exists a source behavior $T$ such that:

- Pruned traces (call/return events removed) are preserved
- Termination/divergence is preserved
- For all metrics $M$, $W_M(T') \leq W_M(T)$
  - Equality holds for most passes (all events preserved)
  - Do not change the metric during a pass (use the assembly metric)
Quantitative compiler correctness

• Given stack size $\beta < 2^{31}$, for all source code $s$, if all the following hold:
  – The compiler produces assembly code $C(s)$ and event metric $M$
  – $s$ does not go wrong in infinite stack space
  – All traces $T$ of $s$ have weight $W_M(T) \leq \beta$
  – Assembly $C(s)$ is run with $\beta$ stack size

• Then:
  – $C(s)$ refines $s$ (I/O events and termination/divergence are preserved)
  – $C(s)$ does not go wrong
  – In particular, $C(s)$ is guaranteed to not stack overflow
Quantitative CompCert

- Function inlining and tailcall recognition underway
- All other passes supported
Quantitative CompCert
CompCert stack management

- CompCert memory model: allocate a fresh stack frame memory block upon function entry
  - No pointer arithmetics across different memory blocks
  - Always succeeds
- Still used for assembly language semantics
  - Requires Pallocframe/Pfreeframe pseudo-instructions to manage stack frame blocks
  - Turned into pointer arithmetics by unverified “pretty-printing” phase
CompCert-generated assembly...

```c
int g(int y);

int f(int x) {
    return g(x-1) - 2;
}
```

- Formal semantics of Pallocframe/Pfreeframe also:
  - stores/loads return address in/from callee's stack frame
    - Uses RA pseudo-register to model caller's return address slot
  - stores/loads back link to caller's stack frame
… after unverified “pretty-printing”

```assembly
f:
  Pallocframe 12, 4
  mov $4(%esp), %edx
  movl (%edx), %eax
  subl $1, %eax
  movl %eax, (%esp)
call g
  subl $2, %eax
  Pfreeframe 12, 4
ret

f:
  subl $8, %esp
  leal $12(%esp), %edx
  movl %edx, 4(%esp)
  mov $4(%esp), %edx
  movl (%edx), %eax
  subl $1, %eax
  movl %eax, (%esp)
call g
  subl $2, %eax
  addl $8, %esp
ret
```

Addresses increase
Stack grows

\[ x = x - 1 \]

\[ y = x - 1 \]
But we can do better and prove it!

```
f:
    subl $8, %esp
    leal $12(%esp), %edx
    movl %edx, 4(%esp)
    mov $4(%esp), %edx
    movl (%edx), %eax
    subl $1, %eax
    movl %eax, (%esp)
    call g
    subl $2, %eax
    addl $8, %esp
    ret
```

```
f:
    subl $4, %esp
    mov $8(%esp), %eax
    subl $1, %eax
    movl %eax, (%esp)
    call g
    subl $2, %eax
    addl $4, %esp
    ret
```
Assembly with finite stack

• Allocate one single stack block at program start
  – Program goes wrong on stack overflow
  – No need for pseudo-instructions
• Merge all stack frames together into the single stack block
  – Requires memory injection proof
Quantitative CompCert
Stack merging

- CompCert Mach to single-stack Mach2 phase
  - Mach already puts arguments into stack
  - Mach no longer stores RA into stack, Mach2 does
  - Mach and Mach2 have same syntax
  - No code transformation: reinterpretation of semantics with single stack

- Mach2 to assembly
  - Implement function entry/exit with stack pointer arithmetics
  - No significant memory changes

- Total changes: 5k LOC (out of CompCert's 90k)
Mach vs. Mach2

- Registers (x86)
  \( r \) := EAX | EBX | ECX | EDX | FP0

- Statements (\( r \)* registers, ofs constant integer)
  \( S ::= M\text{load}(\text{chunk}, \text{raddr}, \text{rres}) \)
  | M\text{store}(\text{chunk}, \text{raddr}, \text{rval})
  | M\text{getstack}(\text{chunk}, \text{ofs}, \text{rres})
  | M\text{setstack}(\text{chunk}, \text{ofs}, \text{rres})
  | M\text{getparam}(\text{chunk}, \text{ofs}, \text{rres})
  | Mcall func  | Mret
  | M\text{goto} label | M\text{label} label: | ...
Mach vs. Mach2

```c
int g(int y);

int f(int x) {
    return g(x-1) - 2;
}

int g {...}

int f {
    Mget param(Mint32, 0, EAX);
    Mop(Osubimm 1, EAX);
    Mset stack(Mint32, 0, EAX);
    Mcall(g);
    Mop(Osubimm 2, EAX);
    Mret
}
```

![Diagram showing differences between Mach and Mach2](image-url)
Overview
Quantitative program logic

- Hoare-like logic
- Assertions have values in \( \{0, 1, 2, \ldots, \infty \} \)
  - Represent available stack space
- \( \{P\} S \{Q\} \) roughly: if P stack space is available before S, then:
  - S does not stack overflow (unless \( P=\infty \)), and
  - for all possible terminating executions of S, Q stack space is available after S
Assertions

- Clight statements $S$, continuations $K$, local state $\theta$
- Global state (“heap” = CompCert memory state) $H$
- Mutable state $\sigma = (\theta, H)$
- Configuration $C = (S, K, \sigma)$
- Assertion $P : C \rightarrow \{0, 1, 2, \ldots, \infty\}$
  - Coq implementation: $C \rightarrow \mathbb{N} \rightarrow \text{Prop}$, represents sets of valid bounds
Selected rules

\[
\frac{\{P\} S_1 \{R\} \quad \{R\} S_2 \{Q\}}{\{P\} S_1; S_2 \{Q\}} \quad (Q:\text{SEQ})
\]

\[
c \geq 0 \quad \frac{\{P\} S \{Q\}}{\{P + c\} S \{Q + c\}} \quad (Q:\text{FRAME})
\]

\[
\frac{P \geq P'}{\{P'\} S \{Q'\}} \quad \frac{Q' \geq Q}{\{P\} S \{Q\}} \quad (Q:\text{CONSEQ})
\]
Selected rules

\[ \frac{\Gamma(f) = (P, Q)}{\Gamma \vdash \{ P + M(f) \} f() \{ Q + M(f) \}} \quad (Q: \text{CALL}) \]
Selected rules

With:

- Global variable addresses $\Delta$
- Loop break
- Return value
- One argument
- Mutable state $(\theta, H)$

those rules become:

\[
\Gamma(f) = (P_f, Q_f) \quad P = \lambda(\theta, H). P_f(\theta(H), H) \quad Q = \lambda(\theta, H). Q_f(\theta(H), H) \\
\Gamma \vdash \{ P + M(f) \} x = f(E) \{(Q + M(f), \perp, \perp)\} \tag{Q:CALL}
\]

\[
\Gamma' = \Gamma, f : (P_f, Q_f) \quad \Sigma(f) = (x, S_f) \\
\Gamma' \vdash \{ P' \} S_f \{ \perp, \perp, Q' \} \quad P' = \lambda(\theta, H). P_f(\theta(x), H) \quad Q' = \lambda(\theta, H). \lambda r. Q_f(r, H) \tag{Q:ABSTRACT}
\]

But we also support:

- Several function arguments
- Auxiliary state
- Stack framing

See paper for more details.
Example with auxiliary state

\[ Z = \log_2(h_\sigma - l_\sigma) \implies M_b \cdot Z \}\]

bsearch(x,l,h) {
  if (h - l <= 1) return l;

  if (Z > 0 \land Z = \log_2(h_\sigma - l_\sigma) \land m_\sigma = \frac{h_\sigma + l_\sigma}{2}) \implies M_b \cdot Z \}

  i = l + (h - l) / 2;

  if (a[m] > x) h = m else l = m;

  [Z - 1 = \log_2(h_\sigma - l_\sigma) \implies M_b \cdot (Z - 1)] + M_b \}

  return bsearch(x, l, h);

  [M_b \cdot (Z - 1)] + M_b \}

  \{ M_b \cdot Z \}
Soundness

- “C consumes at most P stack space” iff for any t, C' such that C \rightarrow_\ast C', and for any metric M, \( W_M(t) \leq P(C, M) \)
- If \{P\} S \{Q\} is derivable, then for any \(\sigma\), (S, Kstop, \(\sigma\)) consumes at most P stack space
  - Stronger soundness: for any K, \(\sigma\)
    if (skip, K, \(\sigma\)) consumes at most Q stack space, then (S, K, \(\sigma\)) consumes at most P stack space
- Logic and soundness: 700 LOC
  Instantiation to Clight: 950 LOC
Accuracy

<table>
<thead>
<tr>
<th>Function Name</th>
<th>Verified Stack Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>recid()</td>
<td>8a bytes</td>
</tr>
<tr>
<td>bsearch(x, lo, hi)</td>
<td>40(1 + log₂(hi - lo)) bytes</td>
</tr>
<tr>
<td>fib(n)</td>
<td>24n bytes</td>
</tr>
<tr>
<td>qsort(a, lo, hi)</td>
<td>48(hi - lo) bytes</td>
</tr>
<tr>
<td>filter_pos(a, sz, lo, hi)</td>
<td>48(hi - lo) bytes</td>
</tr>
<tr>
<td>sum(a, lo, hi)</td>
<td>32(hi - lo) bytes</td>
</tr>
<tr>
<td>fact_sq(n)</td>
<td>40 + 24n² bytes</td>
</tr>
<tr>
<td>filter_find(a, sz, lo, hi)</td>
<td>128 + 48(hi - lo) + 40 log₂(BL) bytes</td>
</tr>
</tbody>
</table>

Table 1. Manually verified stack bounds for C functions.

- **Bound verified manually** using our program logic, then instantiated by CompCert-generated stack frame sizes.

- **Actual stack consumption measured** at run-time thanks to a stack monitor using ptrace (200 lines of C+Perl)

- **4 bytes difference** due to space reserved for RA in the last callee's stack frame

![Graph of fact_sq(x)](image1)

**fact_sq(x)**

![Graph of bsearch(v, lo, hi), x = hi - lo](image2)

**bsearch(v, lo, hi), x = hi - lo**

![Graph of measured stack consumption](image3)

**measured stack consumption 40 + 24 \times 2**

![Graph of measured stack consumption](image4)

**measured stack consumption 40(1 + log₂(x))**
Automatic stack analyzer

- For C code without recursion (e.g. MISRA C), program logic can be automatically applied to derive stack bound
  - 500 lines of Coq

- Instrumented compiler to generate both compiled code and stack bound
  - 400 lines of Coq + 500 Ocaml
Automatic stack analyzer

- Let \( \text{liftO} \{A, B, C : \text{Type}\} \) (\( f: A \to B \to C \) (\( ox: \text{option} \ A \) (\( oy: \text{option} \ B \): \( \text{option} \ C := \ldots \)

- Fixpoint \( B M \Gamma (s: \text{stm}) : \text{option} \ \text{nat} := \)
  match \( s \) with
  | \( \text{scall}_\_ f \_ \Rightarrow \)
    \( \text{liftO} \ \text{plus} \ (\text{Some} \ (M f)) \ (\Gamma f) \)
  | \( \text{sseq} s1 s2 \Rightarrow \)
    \( \text{liftO} \ \text{max} \ (B M \Gamma s1) \ (B M \Gamma s2) \)
  | \( \text{sif} \_ st sf \Rightarrow \)
    \( \text{liftO} \ \text{max} \ (B M \Gamma \Phi st) \ (B M \Gamma sf) \)
  | \( \text{sloop} s \Rightarrow B M \Gamma s \)
  | \_ \Rightarrow \text{Some} \ 0 \)
end.

- Lemma \( \text{sound_B} : \)
  for all \( M \Gamma \) (\( \text{CVALID : valid_bctx} M \Gamma \) s n)
  (\( BS: B M \Gamma s = \text{Some} n \)),
  \( \text{valid_bound} M s n \).

  Proof.
  induction \( s \); intros; ... + apply \( \text{sound_skip} \).
  + apply \( \text{sound_ret} \) with (Q := fun _ \Rightarrow \text{mkassn} 0).
  + apply \( \text{sound_break} \).
  + ... apply \( \text{sound_seq} \) with (Q := fun _ \Rightarrow \text{mkassn} \ (\text{max} \ x \ y) \) ...
  apply \( \text{valid_max}_l \) ... apply \( \text{valid_max}_r \)...
  + case_eq (\( \Gamma f \) ... eapply \( \text{valid_le} \) [ apply \( \text{Le.le_n_Sn} \) ].
    eapply \( \text{sound_consequence} \); [] apply \( \text{sound_call2} \)
    with (C := \( \Gamma \))
    (Pg := fun_pre \( \text{phif} \))
    (Qg := fun_post \( \text{phif} \))
    (L := fun _ _ \Rightarrow \text{True} )].
  + eapply \( \text{CVALID} \); eauto.
  + eapply \( \text{sound_consequence} \); [] apply \( \text{sound_loop} \)
    with (I := fun _ \Rightarrow \text{mkassn} n)
    (Q := fun _ \Rightarrow \text{mkassn} n)
]; unfold \text{mkassn} ; intuition. ...
  eapply IHs; eauto.
Qed.
Automatic stack analyzer: soundness

- Let \( \text{liftO} \) \((A \rightarrow B \rightarrow C) \) \((\text{ox}: \text{option } A) \) \((\text{oy}: \text{option } B)\):
  \[ \text{option } C := \ldots \]

- Fixpoint \( B \ M \Gamma \) \((s: \text{stm})\):
  \[ \text{option } \text{nat} := \]
  \begin{align*}
  \text{match } s \text{ with } \\
  | \text{scall } f \Rightarrow \\
  \quad \text{liftO plus (Some (M f)) (\Gamma f)} \\
  | \text{sseq } s1 \text{ s2} \Rightarrow \\
  \quad \text{liftO max (B M \Gamma s1) (B M \Gamma s2)} \\
  | \text{sif } \text{st} \text{ sf} \Rightarrow \\
  \quad \text{liftO max (B M \Gamma \Phi \text{st}) (B M \Gamma \text{sf})} \\
  | \text{sloop } s \Rightarrow \text{B M \Gamma s} \\
  | _ \Rightarrow \text{Some 0}
  \end{align*}
  \]

- Lemma \( \text{sound_B} \):
  \[ \forall M \Gamma (CVALID: \text{valid_bctx } M \Gamma ) \ s \ n \]
  \[ (BS: \text{B M \Gamma s } = \text{Some } n), \]
  \[ \text{valid_bound } M \Gamma s \ n. \]

- Fixpoint \( \text{bound_of_lvl} \) \((ge \ M \ lvl) \ f := \)
  \[ \text{match lvl with} \]
  \[ | 0 \Rightarrow \text{None} \]
  \[ | S \ lvl' \Rightarrow \]
  \[ \quad \text{match find_func_ } \text{ge } f \text{ with} \]
  \[ \quad | \text{Some bdy } \Rightarrow \]
  \[ \quad \quad \text{B M (bound_of_lvl } \text{ge } M \Gamma \text{vl'}) bdy \]
  \[ \quad | \text{None } \Rightarrow \text{None} \]
  \[ \text{end} \]
  \[ \text{end}. \]

- Theorem \( \text{bound_lvl_sound} \):
  \[ \forall ge \ M \ l, \]
  \[ \text{valid_bctx } M \ (\text{bound_of_lvl } \text{ge } M \Gamma l). \]
  \[ \text{Proof.} \]
  \[ \text{induction } l. \]
  \[ \ldots \text{apply sound_B } \ldots \text{apply IHl } \ldots \]
  \[ \text{Qed.} \]
Automatic stack analyzer: “completeness”

- Fixpoint bound_of_lvl ge M (lvl: nat) f :=
  match lvl with
  | 0 => None
  | S lvl' =>
    match find_func_ge f with
    | Some bdy =>
      B M (bound_of_lvl ge M lvl') bdy
    | None => None
  end
end.

- Theorem bound_of_lvl_complete:
  forall M p (CLOSED: ... p ..) (CG_WELLFOUNDED:forall id fi,
  In (id, Gfun fi) p.(prog_defs) ->
  forall id',
  in_stm id' fi.(fi_body) ->
  id' < id)
  lvl f (LVL: f < lvl)
  fi (FDEF: In (f, Gfun fi) p.(prog_defs)),
  exists n,
  bound_of_lvl (Genv.global env p) M lvl f = Some n.
# Automatic stack bounds

<table>
<thead>
<tr>
<th>File Name / Line Count</th>
<th>Function Name</th>
<th>Verified Stack Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>mibench/net/dijkstra.c (174 LOC)</td>
<td>enqueue</td>
<td>40 bytes</td>
</tr>
<tr>
<td></td>
<td>dequeue</td>
<td>40 bytes</td>
</tr>
<tr>
<td></td>
<td>dijkstra</td>
<td>88 bytes</td>
</tr>
<tr>
<td>mibench/auto/bitcount.c (110 LOC)</td>
<td>bitcount</td>
<td>16 bytes</td>
</tr>
<tr>
<td></td>
<td>bitstring</td>
<td>32 bytes</td>
</tr>
<tr>
<td>mibench/sec/blowfish.c (233 LOC)</td>
<td>BF_encrypt</td>
<td>40 bytes</td>
</tr>
<tr>
<td></td>
<td>BF_options</td>
<td>8 bytes</td>
</tr>
<tr>
<td></td>
<td>BF_ecb_encrypt</td>
<td>80 bytes</td>
</tr>
<tr>
<td>mibench/sec/pgp/md5.c (335 LOC)</td>
<td>MD5Init</td>
<td>16 bytes</td>
</tr>
<tr>
<td></td>
<td>MD5Update</td>
<td>168 bytes</td>
</tr>
<tr>
<td></td>
<td>MD5Final</td>
<td>168 bytes</td>
</tr>
<tr>
<td></td>
<td>MD5Transform</td>
<td>128 bytes</td>
</tr>
<tr>
<td>mibench/tele/fft.c (195 LOC)</td>
<td>IsPowerOfTwo</td>
<td>16 bytes</td>
</tr>
<tr>
<td></td>
<td>NumberOfBitsNeeded</td>
<td>24 bytes</td>
</tr>
<tr>
<td></td>
<td>ReverseBits</td>
<td>24 bytes</td>
</tr>
<tr>
<td></td>
<td>fft_float</td>
<td>160 bytes</td>
</tr>
<tr>
<td>certikos/vmm.c (608 LOC)</td>
<td>palloc</td>
<td>48 bytes</td>
</tr>
<tr>
<td></td>
<td>pfree</td>
<td>40 bytes</td>
</tr>
<tr>
<td></td>
<td>mem_init</td>
<td>72 bytes</td>
</tr>
<tr>
<td></td>
<td>pmap_init</td>
<td>176 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_free</td>
<td>80 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_init</td>
<td>152 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_init_kern</td>
<td>136 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_insert</td>
<td>80 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_read</td>
<td>56 bytes</td>
</tr>
<tr>
<td></td>
<td>pt_resv</td>
<td>120 bytes</td>
</tr>
<tr>
<td></td>
<td>enqueue</td>
<td>48 bytes</td>
</tr>
<tr>
<td></td>
<td>dequeue</td>
<td>48 bytes</td>
</tr>
<tr>
<td></td>
<td>kctxt_new</td>
<td>72 bytes</td>
</tr>
<tr>
<td></td>
<td>sched_init</td>
<td>232 bytes</td>
</tr>
<tr>
<td></td>
<td>tdqueue_init</td>
<td>208 bytes</td>
</tr>
<tr>
<td></td>
<td>thread_init</td>
<td>192 bytes</td>
</tr>
<tr>
<td></td>
<td>thread_spawn</td>
<td>96 bytes</td>
</tr>
<tr>
<td>certikos/proc.c (819 LOC)</td>
<td>main</td>
<td>56 bytes</td>
</tr>
<tr>
<td>compcert/mandelbrot.c (92 LOC)</td>
<td>advance</td>
<td>80 bytes</td>
</tr>
<tr>
<td></td>
<td>energy</td>
<td>56 bytes</td>
</tr>
<tr>
<td></td>
<td>offset_momentum</td>
<td>24 bytes</td>
</tr>
<tr>
<td></td>
<td>setup_bodies</td>
<td>16 bytes</td>
</tr>
<tr>
<td></td>
<td>main</td>
<td>112 bytes</td>
</tr>
</tbody>
</table>

Table 2. Automatically verified stack bounds for C functions.
Conclusion

• Stack overflow need not be enforced by source semantics
  – Stack consumption as add-on to existing operational semantics
• Yet, stack consumption can be verified at the source level and preserved by compilation
• Paves the way for other quantitative properties:
  – Malloc/free heap memory consumption
  – clock cycles, energy...
Thank you!

- Paper (accepted to PLDI 2014, to appear), TR, Coq development and artifact VM: http://cs.yale.edu/~tahina/certikos/stack

- For any questions: tahina.ramananandandro@yale.edu
Function inlining

- void h();
- g() { h(); return 1; }
- f() { int i = g(); return i + 1; }

- Call(f) ::
  - call(g) ::
    - call(h) :: return(h) ::
  - return(g) ::
  - return(f) :: nil

- void h();
- f() { int i = (h(), 1); return i + 1; }

- Call(f) ::
  - call(h) :: return(h) ::
  - return(f) :: nil

- Events are removed *in matching pairs*
Function inlining

\[(\text{call}(f) \cdot T, \theta) \rightsquigarrow (T, \text{call}(f) \cdot \theta)\]
\[(\text{ret}(f) \cdot T, \text{call}(f) \cdot \theta) \rightsquigarrow (T, \theta)\]

With \(\theta\) finite and only containing call events

Coinductively:

\(\epsilon \sqsubseteq_\theta \epsilon\)
\(e \cdot T' \sqsubseteq_\theta t \cdot e \cdot T\)
\((t \cdot e \cdot T, \theta) \rightsquigarrow^* (e \cdot T, \theta')\)
\(\epsilon \sqsubseteq_\theta T\)
\(\text{if } T' \sqsubseteq_\theta T\)

\(\text{if } \epsilon \sqsubseteq_\theta \epsilon'\)

\(\text{if } e \sqsubseteq_\theta T'\)

\(\text{and } (T, \theta) \rightsquigarrow^+ (T', \theta')\)

- If \(T' \sqsubseteq_\theta T\), then
  - for \(t'\) finite prefix of \(T'\)
    - there is \(t\) finite prefix of \(T\)
      - such that
        \(V_M(t') - V_M(\theta) \leq V_M(t)\)
  - So,
    \(W_M(T') - W_M(\theta) \leq W_M(T)\)

- Thus, it suffices to prove that for any \(T'\) of the target,
  there is \(T\) of the source such that \(T' \sqsubseteq_\epsilon T\)
Tailcall recognition

```c
int h();
int g(x)
{ return h(x+1); }
int f(x)
{ return g(x+2); }
```

- Call(f) ::
  - call(g) ::
    - call(h) :: return(h) ::
      - return(g) ::
        - return(f) :: nil

- Caller produces return event *before* transferring to tail-callee

- call(f) :: return(f) :: call(g) :: return(g) :: call(h) :: return(h)
Tailcall recognition

With $\theta$ finite and only containing return events

Coinductively:

\[
\begin{align*}
\epsilon & \sqsubseteq_{\theta} \epsilon \\
e \cdot T' & \sqsubseteq_{\theta} e \cdot T & \text{if } T' \sqsubseteq_{\theta} T \\
\text{ret}(f) \cdot e & \sqsubseteq_{\theta} e & \text{if } T' \sqsubseteq_{\theta} T \\
\text{ret}(f) \cdot e \cdot T' & \sqsubseteq_{\theta} e \cdot T & \text{if } T' \sqsubseteq_{\text{ret}(f) \cdot \theta} T \\
T' & \sqsubseteq_{\text{ret}(f) \cdot \theta} \text{ret}(f) \cdot T & \text{if } T' \sqsubseteq_{\theta} T
\end{align*}
\]

- If $T' \sqsubseteq_{\theta} T$, then

  - for $t'$ finite prefix of $T'$ there is $t$ finite prefix of $T$ such that $V_M(t') + V_M(\theta) \leq V_M(t)$
  - So, $W_M(T') + W_M(\theta) \leq W_M(T)$

- Thus, it suffices to prove that for any $T'$ of the target, there is $T$ of the source such that $T' \sqsubseteq_{\epsilon} T$
Function inlining and tailcall recognition

- Need to modify simulation diagrams to take special refinement relations into account
- Proof in progress
Mach configuration

- **Continuations**
  \[ K ::= \text{Knil} | \text{Kcons}(SP, f, code, \text{RA}, K) \]

- **Configurations**
  \[ C ::= \text{State}(\text{mem}, \text{rset}, SP, f, code, K) \]
  \[ | \text{Callstate}(\text{mem}, \text{rset}, f, K) \]
  \[ | \text{Returnstate}(\text{mem}, \text{rset}, K) \]

- Callstate/Returnstate correspond to CompCert assembly Pallocframe/Pfreeframe pseudos
Mach2 configuration

- **Continuations**
  \[ K ::= \text{Knil} \mid \text{Kcons}(f, \text{code}, \text{RA, } K) \]

- **Configurations**
  \[ C ::= \text{State}(\text{mem, rset, SP, f, code, K}) \]
  \[ \mid \text{Callstate}(\text{mem, rset, f, K}) \]
  \[ \mid \text{Returnstate}(\text{mem, rset, K}) \]

- Callstate/Returnstate do not modify the stack
Thank you!

- Paper (accepted to PLDI 2014, to appear), TR, Coq development and artifact VM: http://cs.yale.edu/~tahina/certikos/stack

- For any questions: tahina.ramananandro@yale.edu