Hybrid Dynamic Data Race Detection

Robert O’Callahan
Jong-Deok Choi
IBM T.J. Watson Research Center

ACM SIGPLAN 2003 Symposium on Principles and Practice of Parallel Programming (PPoPP’03)
San Diego, CA, June 11–13

Presents a method for dynamic detection of data races
Synergistically combines lockset-based detection and happens before-based detection
– Fewer false positives than just lockset-based detection
– Fewer false negatives than just happens before-based detection
– Little extra computational expense over lockset-based detection
• Provides more diagnostic information about potential race
  – Including stack traces for both threads involved in the conflicting events
• Improves scalability of race detection
  – Eliminates need for whole-program static analysis to justify optimizations
  – Instead:
    ◦ Perform dynamic optimizations: thread-locality check and over-sized lockset check to identify redundant events
    ◦ Run program twice with different instrumentation in each run

Key contributions

Program execution model

Scenario: Their race detector observes a stream of events generated by instrumentation inserted into the program and serialized via locks inside the detector.

This scenario justifies:
A program is an abstract machine that outputs a sequence of events to the detector.
The components of an event are abstracted by:
• A set of memory locations, $M$
• A set of locks, $L$
• A set of threads, $T$
• A set of message IDs, $G$
• A set of access types, $A = \{ \text{RD}, \text{WR} \}$

Notational conventions:

The events generated by a program are abstracted by:
• $\text{MEM}(m, a, t)$: thread $t$ performed an access of type $a$ to location $m$.
• $\text{ACQ}(l, t)$: thread $t$ acquired lock $l$ (and did not already hold it).
• $\text{REL}(l, t)$: thread $t$ released lock $l$ (and no longer holds it).
• $\text{SND}(g, t)$: thread $t$ sent the message $g$.
• $\text{RCV}(g, t)$: thread $t$ received the message $g$.

Latter two types are observed exclusively by the happens-before detector; lockset detector observes only memory accesses and lock events.
Example: events generated (for lockset algorithm)

```java
// MAIN THREAD
class Main {
    int gflag;
    CThr cthr;
    void execute() {
        gflag = 1;
        // MEM(main.gflag,WT,MAIN)
        cthr = new CThr(this);
        // MEM(cthr.main,WT,MAIN)
        cthr.start();
        // MEM(cthr.main,RD,MAIN)
        ...
    }

    synchronized (this) {
        // ACQ(main,MAIN)
        L: cthr.interrupt();
        // MEM(cthr.main,WT,MAIN)
    }
    // REL(main,MAIN)
}
```

Lockset-based detection: preliminaries

Given an event sequence \( <e_i> \), the locks before step \( i \) held by a thread \( t \), \( L(t) \), is defined as:

\[
L_i(t) = \{ l | (\exists j \land j < i \land e_j = ACQ(l, t) \land \forall k, j < k < i \Rightarrow e_k \neq REL(l, t) ) \}
\]

The lockset hypothesis: whenever two threads access the same location and one of the accesses is a WT, the accesses are performed holding a common lock.

\[
PossLSR(i, j) = e_i = MEM(m, a, t) \land e_j = MEM(m, a, t') \land t \neq t' \land (a = WT \lor a' = WT) \land L_i(t) \cap L_j(t') = \emptyset
\]

Lockset-based detection: example
Lockset-based detection: false positives

Violations of the lockset hypothesis are not necessarily synchronization errors

- In the example, the reported PossLSR on $m.gflag$ is not a true race since $gflag = 1$ necessarily happens before `cthr.start()` in Main, while the latter necessarily happens before evaluation of $m.gflag == 1$ in CThr.
- More generally, programs may use message passing for synchronization instead of locks

```
// THREAD A
BigObj big = new BigObj();
big.init(...);
... pool.recycle(big);
```

```
// THREAD B
BigObj big = pool.removeOne();
big.init(...);
... pool.recycle(big);
```

Since threads access `big` without any locks, a potential race will be reported.

Happens before-based detection: preliminaries

Key idea: $e$, happens before $e'$, if communication between threads could allow information to be transferred from $e$ to $e'$.

Message send and receive events are introduced to capture this relation.

Formally, given an event sequence $<e_i>_{i=1,...}$, the happens-before relation, $\rightarrow$, is the smallest relation satisfying

- if $\text{Thread}(e) = \text{Thread}(e')$ and $i < j$, then $i \rightarrow j$
- if $e_i = \text{SND}(g,t)$ \& $e_j = \text{RCV}(g,t')$, then $i \rightarrow j$
- if $i \rightarrow j \& j \rightarrow k$, then $i \rightarrow k$

It is defined over indices because the same event may occur many times in an execution and each occurrence has its own happens-before relationships.

Happens before-based detection: messages

Key idea: Want a message to be sent from $t_1$ to $t_2$ whenever $t_1$ takes an action that later affects $t_2$.

Capture explicit synchronization in Java program, e.g,

- When $t_1$ calls $t_2.start()$, generate a $\text{SND}(\varphi, t_1)$ and, when $t_1$ first executes in `run()`, generate a $\text{RCV}(\varphi, t_2)$, using a fresh $\varphi$.
- When $t_2$ calls `obj.notify()`, generate a $\text{SND}(\varphi, t_2)$ and also $\text{RCV}(\varphi, t_2)'s$, for all $t_2$ waiting on `obj`, using a fresh $\varphi$.

Capture interactions arising from shared memory accesses

- After `read(m, wr, l)` generate a $\text{SND}(\varphi, l)$ using a fresh $\varphi$.
- After `read(m, rd, l)` generate a $\text{SND}(\varphi, l)$ using the $\varphi$ sent by the most recent writer to $m$. (I'm not generating this if most recent writer is $l$. I think that's OK.)

Capture interactions arising from locking

- After `rel(l, l)` generate a $\text{SND}(\varphi, l)$ using a fresh $\varphi$.
- After (all but the first) `acq(l, l)` generate a $\text{SND}(\varphi, l)$ using the $\varphi$ sent by the most recent `rel(l, l)`. 
Happens before-based detection: messages

```java
// M THREAD (MAIN)
gflag = 1;
// MEM(m,gflag,WT,M)
// SND(gflag1,M)
cthr = new CThr(this);
// MEM(m,cthr,WT,M)
// SND(cthr1,M)
cthr.start();
// MEM(m,cthr,RD,M)
// SND(start,M)

synchronized {this}{
    // ACQ(m,M)
    this.m = main; not checking
    // RCV(start,C)
    if (m,gflag == 1)
        // MEM(m,cthr,WT,C)
        // RCV(gflag1,C)

    main.cthr = null;
    // MEM(m,cthr,WT,C)
    // SND(cthr2,C)
}
```

```
L. Dillon, CSE 914, Fall 2008 13
```

Happens before-based detection

A potential race occurs if two events by different threads access the same location, at least one access is a write, and the accesses are not causally ordered by the happens-before relation.

Formally:

\[
\text{Poss_{HBR}(i,j) = } \begin{cases} 
    e_i = \text{MEM}(m,a,t) \land e_j = \text{MEM}(m,a,t) \land t_i \neq t_j \land \\
    (a_i = \text{WT} \lor a_j = \text{WT}) \land L_i(t) \cap L_j(t) = \emptyset \land i \neq j \land j \neq i
\end{cases}
\]

The happens-before relation is computed on-line using vector-clocks

Happens before-based detection: vector clocks

Each thread maintains a vector clock indexed by thread IDs:
- the entry for \( t' \) indicates the last event in \( t \) that could have influenced \( t \).

Each \( g \) carries a vector clock \( V(g) \)
- the vector clock of the sending thread at the time \( g \) is sent.

Given an event sequence \( <e_i>_{i=1,2,\ldots} \), the vector clock \( V(t) \) of \( t \) at the completion of \( e \) is defined inductively:
- Initial vector clock: \( V_i(t_i) = 1 \) and \( V_i(t_i') = 0 \), for \( i \neq t' \)
- Subsequently:
  - If \( e_i = \text{SND}(g,t) \), then \( V_i(t) = V_i(t') + 1 \)
  - If \( e_i = \text{RCV}(g,t) \) and \( t \neq t' \), then \( V_i(t') = \max(V_i(t'), V_i(t)) \)
  - Otherwise, \( V_i(t) = V_i(t') \)

Additionally, \( V(g) = V(t) \), where \( e_i = \text{SND}(g,t) \)

```
Thread 1 Thread 2 Thread 3

\[
\begin{array}{ccc}
V_1 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_2 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_3 & (2,0,0) & (0,2,0) & (0,1,1) \\
V_4 & (0,2,0) & (0,1,1) & V(g1) \\
V_5 & (2,0,0) & (2,0,0) & (0,1,1) \\
V_6 & (2,0,0) & (1,2,0) & (0,1,2) \\
V_7 & (2,0,0) & (1,2,0) & (0,1,2) \\
\end{array}
\]

\( i \rightarrow j \) iff \( i < j \) and \( V(t_i)(t_j) = \min(V(t_i)(t_j)) \) where \( t_i = \text{Thread}(e_i) \) and \( t_j = \text{Thread}(e_j) \)
```

```
L. Dillon, CSE 914, Fall 2008 14
```

Happens before-based detection: vector clocks

```
Thread 1 Thread 2 Thread 3

\[
\begin{array}{ccc}
V_1 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_2 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_3 & (2,0,0) & (0,2,0) & (0,1,1) \\
V_4 & (0,2,0) & (0,1,1) & V(g1) \\
V_5 & (2,0,0) & (2,0,0) & (0,1,1) \\
V_6 & (2,0,0) & (1,2,0) & (0,1,2) \\
V_7 & (2,0,0) & (1,2,0) & (0,1,2) \\
\end{array}
\]

\( i \rightarrow j \) iff \( i < j \) and \( V(t_i)(t_j) = \min(V(t_i)(t_j)) \) where \( t_i = \text{Thread}(e_i) \) and \( t_j = \text{Thread}(e_j) \)
```

```
L. Dillon, CSE 914, Fall 2008 15
```

Happens before-based detection: vector clocks

```
Thread 1 Thread 2 Thread 3

\[
\begin{array}{ccc}
V_1 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_2 & (1,0,0) & (0,1,0) & (0,0,1) \\
V_3 & (2,0,0) & (0,2,0) & (0,1,1) \\
V_4 & (0,2,0) & (0,1,1) & V(g1) \\
V_5 & (2,0,0) & (2,0,0) & (0,1,1) \\
V_6 & (2,0,0) & (1,2,0) & (0,1,2) \\
V_7 & (2,0,0) & (1,2,0) & (0,1,2) \\
\end{array}
\]

\( i \rightarrow j \) iff \( i < j \) and \( V(t_i)(t_j) = \min(V(t_i)(t_j)) \) where \( t_i = \text{Thread}(e_i) \) and \( t_j = \text{Thread}(e_j) \)
```

```
L. Dillon, CSE 914, Fall 2008 16
```
Relation between HBD and LSD

If a full HB detector reports a race, then there are thread schedules in which the two events occur consecutively in time in either order.

The potential races reported by a full HB detector are a subset of those reported by a LS detector.

A full HB detector is subject to more false negatives than a LS detector—i.e., it may fail to detect real races that would be detected by a LS detector.

• Core problem: not all happens-before relationships correspond to true synchronization.

Example false negative introduced by HB

```c
// THREAD A
// THREAD B
global = 7; // *
// MEM(global,WT,A)
... synchronize (1) {
    // ACQ(1,A)
    ctr ++;
    ...
} // REL(1,A)
// SND(l1,A)

// THREAD B
synchronize (1) {
    // ACQ(1,B)
    ctr++
    ...
} // REL(1,B)
// SND(l2,B)
int x = global; // **
// MEM(global,RO,B)
```

Here, the happens-before relation masks the true race between the accesses at * and **

Hybrid race detection

Add limited HB checking to a LS detector
• Generate messages for start(), join(), wait(), notify()
• No messages generated for memory access
• No messages generated for locking

Relation computed using limited HB information is a subset of the full HB relation; use it instead of → to compute PossHybridRace(i,j).

Observations
• Significantly fewer messages means significantly fewer vector clock updates, which are expensive; HB checks are cheap
• Store thread’s local time (not full vector clock) for all stored tuples:
  - (a,t,L,v), for e_i = MEM(m,a,t), L = L(t), v = V(t) (t)
• Maintain full vector clocks for only active threads and pending messages

Hybrid race detection: the algorithm

Simple version: for each m ∈ M, maintains
• a set S_m of tuples (a, t, L, v) representing an access of type a to location m by thread t at local time v while holding lockset L.
• a flag checking_m indicating whether accesses to m are being checked.

Given an event sequence <e_i>_{i=1..},

For each m ∈ M do: initialize S_m = ∅ ∧ checking_m = true

For each e_i ∈ <e_i>_{i=1..} do:

if (e_i = MEM(m,a,t) ∧ checking_m) {
    if ((a, t, L(t), V(t) (t)) ∈ S_m) {          // e_i is redundant, ignore it
    }
    else if (∃(a', t', L',v') ∈ S_m.         // e_i is a race
        (a = WT ∨ a' = WT) ∧ t ≠ t' ∧ L(t) ∨ L'(t') ∧ V(t) (t) < v')
        report a race
        checking_m = false
    } else {                                    // remember the new tuple
        S_m = S_m ∪ {(a, t, L(t), V(t) (t))}
    }
}
Hybrid detection: example

```
// N THREAD (NAI)
// C THREAD (CTHR)
// m.gflag m.cthr V(M) V(C)

gflag = 1;
// MEM(m.gflag,WT,M)
(WT, 0, M, 1) 1.0 0.1

cthr = new CThr...
// MEM(m.cthr,WT,M)

cthr.start();
// MEM(m.cthr,RD,M)
// SEND(start,M)

this.m = m;
// RCV(start,M)
// MEM(c.m,WT,C)
(RD, 0, C, 1)

if (m.gflag == 1)
// MEM(m.gflag, RD, C)
(RD, 0, C, 1)

synchronized (this)
// ACQ(m,H)

m.cthr = null;
// MEM(m.cthr,WT,C)
```

Hybrid race detection: optimizations

- **Heuristics to identify redundant events.**
  - **Lockset-subset condition:** If the thread, access type, and timestamp of a previous event $e_i$ match those of a new event $e_j$ and if $e_i$’s lockset is a subset of $e_j$’s lockset, then is $e_j$ redundant.
  
  - **Oversized lockset condition:**
    ⋆ Based on two observations (see paper for details)
    1. The number of threads held by any thread at any time tends to be small.
    2. As the algo collects more non-racing, non-redundant accesses to a location, for a new access to be non-redundant, the size of the lockset of future racing events must be large.
    ⋆ Helps reduce overhead of checking accesses of "read only" data, which threads tend to access while holding various unrelated locks.

Hybrid race detection: experimental results

"... our tool achieves very precise results with reasonable slowdown"

See the paper for details. :-)

Hybrid race detection: two-phase mode selection

- **Two-phase mode selection**
  - Run detector first in "simple" mode to identify data that potentially incur races.
    ⋆ In this mode, detector performs LS-based detection equivalent to Eraser
    ⋆ A thread-locality check is performed by extending each field with an owning thread, which is set to the first thread that accesses the field; access by a non-owning thread triggers logging of accesses.
  - Rerun the detector in "detailed" mode, instrumenting accesses to only "race-prone" fields as detected in simple mode.