

Trace-Based Run-Time Analysis of Message-Passing Go Programs

Martin Sulzmann and Kai Stadtmüller

Faculty of Computer Science and Business Information Systems
Karlsruhe University of Applied Sciences
Moltkestrasse 30, 76133 Karlsruhe, Germany
`martin.sulzmann@hs-karlsruhe.de`
`kai.stadtmueller@live.de`

Abstract. We consider the task of analyzing message-passing programs by observing their run-time behavior. We introduce a purely library-based instrumentation method to trace communication events during execution. A model of the dependencies among events can be constructed to identify potential bugs. Compared to the vector clock method, our approach is much simpler and has in general a significant lower run-time overhead. A further advantage is that we also trace events that could not commit. Thus, we can infer more alternative communications. This provides the user with additional information to identify potential bugs. We have fully implemented our approach in the Go programming language and provide a number of examples to substantiate our claims.

1 Introduction

We consider run-time analysis of programs that employ message-passing. Specifically, we consider the Go programming language [4] which integrates message-passing in the style of Communicating Sequential Processes (CSP) [6] into a C style language. We assume the program is instrumented to trace communication events that took place during program execution. Our objective is to analyze program traces to assist the user in identifying potential concurrency bugs.

Motivating Example In Listing 1.1 we find a Go program implementing a system of newsreaders. The `main` function creates two synchronous channels, one for each news agency. Go supports (a limited form of) type inference and therefore no type annotations are required. Next, we create one thread per news agency via the keyword `go`. Each news agency transmits news over its own channel. In Go, we write `ch <- "REUTERS"` to send value "REUTERS" via channel `ch`. We write `<-ch` to receive a value via channel `ch`. As we assume synchronous channels, both operations block and only unblock once a sender finds a matching receiver. We find two newsreader instances. Each newsreader creates two helper threads that wait for news to arrive and transfer any news that has arrived to a common channel. The intention is that the newsreader wishes to receive *any* news whether it be from Reuters or Bloomberg. However, there is a subtle bug (to be explained shortly).

```

func reuters(ch chan string) { ch <- "REUTERS" } // r!
func bloomberg(ch chan string) { ch <- "BLOOMBERG" } // b!

func newsReader(rCh chan string, bCh chan string) {
    ch := make(chan string)
    go func() { ch <- (<-rCh) }()           // r?; ch!
    go func() { ch <- (<-bCh) }()           // b?; ch!
    x := <-ch                               // ch?
}

func main() {
    reutersCh := make(chan string)
    bloombergCh := make(chan string)
    go reuters(reutersCh)
    go bloomberg(bloombergCh)
    go newsReader(reutersCh, bloombergCh) // N1
    newsReader(reutersCh, bloombergCh)    // N2
}

```

Listing 1.1. Message passing in Go

Trace-Based Run-Time Verification We only consider finite program runs and therefore each of the news agencies supplies only a finite number of news (exactly one in our case) and then terminates. During program execution, we trace communication events, e.g. send and receive, that took place. Due to concurrency, a bug may not manifest itself because a certain ‘bad’ schedule is rarely taken in practice.

Here is a possible trace resulting from a ‘good’ program run.

```
r!; N1.r?; N1.ch!; N1.ch?; b!; N2.b?; N2.ch!; N2.ch?
```

We write **r!** to denote that a send event via the Reuters channel took place. As there are two instances of the **newsReader** function, we write **N1.r?** to denote that a receive event via the local channel took place in case of the first **newsReader** call. From the trace we can conclude that the Reuters news was consumed by the first newsreader and the Bloomberg news by the second newsreader.

Here is a trace resulting from a bad program run.

```
r!; b!; N1.r?; N1.b?; N1.ch!; N1.ch?; DEADLOCK
```

The helper thread of the first newsreader receives the Reuters *and* the Bloomberg news. However, only one of these messages will actually be read (consumed). This is the bug! Hence, the second newsreader gets stuck and we encounter a deadlock. The issue is that such a bad program run may rarely show up. So, the question is how can we assist the user based on the trace information resulting from a good program run? How can we infer that alternative schedules and communications may exist?

Event Order via Vector Clock Method A well-established approach is to derive a partial order among events. This is usually achieved via a vector of (logical) clocks. The vector clock method was independently developed by Fidge [1] and Mattern [8]. For the above good program run, we obtain the following partial order among events.

```

r! < N1.r?          b! < N2.b?
N1.r? < N1.ch!      N2.b? < N2.ch!  (1)
N1.ch! < N1.ch?     N2.ch! < N2.ch? (2)

```

For example, (1) arises because `N2.ch!` happens (sequentially) after `N2.b?` For synchronous send/receive, we assume that receive happens after send. See (2). Based on the partial order, we can conclude that alternative schedules are possible. For example, `b!` could take place before `r!`. However, it is not clear how to infer alternative communications. Recall that the issue is that one of the newsreaders may consume both news messages. Our proposed method is able to clearly identify this issue and has the advantage to require a much simpler instrumentation. We discuss these points shortly. First, we take a closer look at the details of instrumentation for the vector clock method.

Vector clocks are a refinement of Lamport's time stamps [7]. Each thread maintains a vector of (logical) clocks of all participating partner threads. For each communication step, we advance and synchronize clocks. In pseudo code, the vector clock instrumentation for event `sndR`.

```

vc[reutersThread]++
ch <- ("REUTERS", vc, vcCh)
vc' := max(vc, <-vcCh)

```

We assume that `vc` holds the vector clock. The clock of the Reuters thread is incremented. Besides the original value, we transmit the sender's vector clock and a helper channel `vcCh`. For convenience, we use tuple notation. The sender's vector clock is updated by building the maximum among all entries of its own vector clock and the vector clock of the receiving party. The same vector clock update is carried out on the receiver side.

Our Method We propose a much simpler instrumentation and tracing method to obtain a partial order among events. Instead of a vector clock, each thread traces the events that might happen and have happened. We refer to them as pre and post events. In pseudo code, our instrumentation for `sndR` looks like follows.

```

pre(hash(ch), "!")
ch <- ("REUTERS", threadId)
post(hash(ch), "!")

```

The bang symbol (!) indicates a send operation. Function `hash` builds a hash index of channel names. The sender transmits its thread id number to the receiver. This is the only intra-thread overhead. No extra communication link is necessary.

Here are the traces for individual threads resulting from the above good program run.

```
R:          pre(r!); post(r!)
N1_helper1: pre(r?); post(R#r?); pre(ch1!); post(ch1!)
N1_helper2: pre(b?)
N1:         pre(ch1?); post(N1_helper1#ch1?)
B:          pre(b!); post(b!)
N2_helper1: pre(r?)
N2_helper2: pre(b?); post(B#b?); pre(ch2!); post(ch2!)
N2:         pre(ch2?); post(N2_helper2#ch2?)
```

We write `pre(r!)` to indicate that a send via the Reuters channel might happen. We write `post(R#r?)` to indicate that a receive has happened via thread R. The partial order among events is obtained by a simple post-processing phase where we linearly scan through traces. For example, within a trace there is a strict order and therefore

```
N2_helper2: pre(b?); post(B#b?); pre(ch2!); post(ch2!)
```

implies `N2.b? < N2.ch!`. Across threads we check for matching pre/post events. Hence,

```
R:          pre(r!); post(r!)
N1_helper1: pre(r?); post(R#r?); ...
```

implies `r! < N1.r?`. So, we obtain the same (partial order) information as the vector clock approach but with less overhead.

The reduction in terms of tracing overhead compared to the vector clock method is rather drastic assuming a library-based tracing scheme with no access to the Go run-time system. For each communication event we must exchange vector clocks, i.e. n additional (time stamp) values need to be transmitted where n is the number of threads. Besides extra data to be transmitted, we also require an extra communication link because the sender requires the receivers vector clock. In contrast, our method incurs a constant tracing overhead. Each sender transmits in addition its thread id. No extra communication link is necessary. This results in *much* less run-time overhead as we will see later.

The vector clock tracing method can be improved assuming we extend the Go run-time system. For example, by maintaining a per-thread vector clock and having the run-time system carrying out the exchange of vector clocks for each send/receive communication. There is still the $O(n)$ space overhead. Our method does not require any extension of the Go run-time system to be efficient and therefore is also applicable to other languages that offer similar features as found in Go.

A further advantage of our method is that we also trace (via pre) events that could not commit (post is missing). Thus, we can easily infer alternative communications. For example, for R: `pre(r!); ...` there is the alternative match

`N2_helper1: pre(r?)`. Hence, instead of $r! < N1.r?$ also $r! < N2.r?$ is possible. This indicates that one newsreader may consume both news message. The vector clock method, only traces events that could commit, post events in our notation. Hence, the above alternative communication could not be derived.

Contributions Compared to earlier works based on the vector clock method, we propose a much more light-weight and more informative instrumentation and tracing scheme. Specifically, we make the following contributions:

- We give a precise account of our run-time tracing method (Section 3) for message-passing as found in the Go programming language (Section 2) where for space reasons we only formalize the case of synchronous channels and selective communications.
- A simple analysis of the resulting traces allows us to detect alternative schedules and communications (Section 4). For efficiency reasons, we employ a directed dependency graph to represent happens-before relations (Section 4.1).
- We show that vector clocks can be easily recovered based on our tracing method (Section 5). We also discuss the pros and cons of both methods for analysis purposes.
- Our tracing method can be implemented efficiently as a library. We have fully implemented the approach supporting all Go language features dealing with message-passing such as buffered channels, select with default or timeout and closing of channels (Section 6).
- We provide experimental results measuring the often significantly lower overhead of our method compared to the vector clock method assuming based methods are implemented as libraries (Section 6.2).

The online version of this paper contains an appendix with further details.¹

2 Message-Passing Go

Syntax For brevity, we consider a much simplified fragment of the Go programming language. We only cover straight-line code, i.e. omitting procedures, if-then-else etc. This is not an onerous restriction as we only consider finite program runs. Hence, any (finite) program run can be represented as a program consisting of straight-line code only.

Definition 1 (Program Syntax).

x, y, \dots	<i>Variables, Channel Names</i>
i, j, \dots	<i>Integers</i>
$b ::= x \mid i \mid \text{hash}(x) \mid \text{head}(b) \mid \text{last}(b) \mid bs \mid \text{tid}$	<i>Expressions</i>
$bs ::= [] \mid b : bs$	
$e, f ::= x \leftarrow b \mid y := \leftarrow x$	<i>Transmit/Receive</i>
$c ::= y := b \mid y := \text{makeChan} \mid \text{go } p \mid \text{select } [e_i \Rightarrow p_i]_{i \in I}$	<i>Commands</i>
$p, q, r ::= [] \mid c : p$	<i>Program</i>

¹ <https://arxiv.org/abs/1709.01588>

For our purposes, values are integers or lists (slices in Go terminology). For lists we follow Haskell style notation and write $b : bs$ to refer to a list with head element b and tail bs . We can access the head and last element in a list via primitives `head` and `last`. We often write $[b_1, \dots, b_n]$ as a shorthand $b_1 : \dots : []$. Primitive `tid` yields the thread id number of the current thread. We assume that the main thread always has thread id number 1 and new thread id numbers are generated in increasing order. Primitive `hash()` yields a unique hash index for each variable name. Both primitives show up in our instrumentation.

A program is a sequence of commands where commands are stored in a list. Primitive `makeChan` creates a new synchronous channel. Primitive `go` creates a new go routine (thread). For send and receive over a channel we follow Go notation. We assume that a receive is always tied to an assignment. For assignment we use symbol $:=$ to avoid confusion with the mathematical equality symbol $=$. In Go, symbol $:=$ declares a new variable with some initial value. We also use $:=$ to overwrite the value of existing variables. As a message passing command we only support selective communication via `select`. Thus, we can fix the bug in our newsreader example.

```
func newsReaderFixed(rCh chan string, bCh chan string) {
    ch := make(chan string)
    select {
        case x := <-rCh:
        case x := <-bCh:
    }
}
```

The `select` statement guarantees that at most one news message will be consumed and blocks if no news are available. In our simplified language, we assume that the $x \leftarrow b$ command is a shorthand for `select [x ← b ⇒ []]`. For space reasons, we omit buffered channels, `select` paired with a default/timeout case and closing of channels. All three features are fully supported by our implementation.

Trace-Based Semantics The semantics of programs is defined via a small-step operational semantics. The semantics keeps track of the trace of channel-based communications that took place. This allows us to relate the traces obtained by our instrumentation with the actual run-time traces.

We support multi-threading via a reduction relation

$$(S, [i_1 \# p_1, \dots, i_n \# p_n]) \xRightarrow{T} (S', [j_1 \# q_1, \dots, j_n \# q_n]).$$

We write $i \# p$ to denote a program p that runs in its own thread with thread id i . We use lists to store the set of program threads. The state of program variables, before and after execution, is recorded in S and S' . We assume that threads share the same state. Program trace T records the sequence of communications that took place during execution. We write $x!$ to denote a send operation on channel x and $x?$ to denote a receiver operation on channel x . The semantics of expressions is defined in terms a big-step semantics. We employ a reduction

relation $(i, S) \vdash b \Downarrow v$ where S is the current state, b the expression and v the result of evaluating b . The formal details follow.

Definition 2 (State).

$$\begin{array}{ll} v ::= x \mid i \mid [] \mid vs & \text{Values} \\ vs ::= [] \mid v : vs & \\ s ::= v \mid \text{Chan} & \text{Storables} \\ S ::= () \mid (x \mapsto s) \mid S \triangleleft S & \text{State} \end{array}$$

A state S is either empty, a mapping, or an override of two states. Each state maps variables to storables. A storable is either a plain value or a channel. Variable names may appear as values. In an actual implementation, we would identify the variable name by a unique hash index. We assume that mappings in the right operand of the map override operator \triangleleft take precedence. They overwrite any mappings in the left operand. That is, $(x \mapsto v_1) \triangleleft (x \mapsto v_2) = (x \mapsto v_2)$.

Definition 3 (Expression Semantics $(i, S) \vdash b \Downarrow v$).

$$\begin{array}{c} \frac{S(x) = v}{(i, S) \vdash x \Downarrow v} \quad (i, S) \vdash j \Downarrow j \quad (i, S) \vdash [] \Downarrow [] \quad \frac{(i, S) \vdash b \Downarrow v \quad (i, S) \vdash bs \Downarrow vs}{(i, S) \vdash b : bs \Downarrow v : vs} \\ \frac{(i, S) \vdash b \Downarrow v : vs}{(i, S) \vdash \text{head}(b) \Downarrow v} \quad \frac{(i, S) \vdash b \Downarrow [v_1, \dots, v_n]}{(i, S) \vdash \text{last}(b) \Downarrow v_n} \quad (i, S) \vdash \text{tid} \Downarrow i \quad (i, S) \vdash \text{hash}(x) \Downarrow x \end{array}$$

Definition 4 (Program Execution $(S, P) \xRightarrow{T} (S', Q)$).

$$\begin{array}{ll} i\sharp p & \text{Single program thread} \\ P, Q ::= [] \mid i\sharp p : P & \text{Program threads} \\ t & ::= i\sharp x! \mid i \leftarrow j\sharp x? \text{ Send and receive event} \\ T & ::= [] \mid t : T \text{ Trace} \end{array}$$

We write $(S, P) \Rightarrow (S', Q)$ as a shorthand for $(S, P) \xRightarrow{} (S', Q)$.

Definition 5 (Single Step).

$$\begin{array}{c} (\text{Terminate}) \quad (S, i\sharp [] : P) \Rightarrow (S, P) \\ (\text{Assign}) \quad \frac{(i, S) \vdash b \Downarrow v \quad S' = S \triangleleft (y \mapsto v)}{(S, i\sharp (y := b : p) : P) \Rightarrow (S', i\sharp p : P)} \\ (\text{MakeChan}) \quad \frac{S' = S \triangleleft (y \mapsto \text{Chan})}{(S, i\sharp (y := \text{makeChan} : p) : P) \Rightarrow (S', i\sharp p : P)} \end{array}$$

Definition 6 (Multi-Threading and Synchronous Message-Passing).

$$\begin{aligned}
(Go) \quad & \frac{i \notin \{i_1, \dots, i_n\}}{(S, i_1 \# (\text{go } p : p_1) : P) \Rightarrow (S, i_1 \# p : i_1 \# p_1 : P)} \\
(Sync) \quad & \frac{\begin{array}{c} \exists l \in J, m \in K. e_l = x \leftarrow b \quad f_m = y := \leftarrow x \quad S(x) = Chan \\ (i_1, S) \vdash b \Downarrow v \quad S' = S \triangleleft (y \mapsto v) \end{array}}{(S, i_1 \# (\text{select } [e_j \Rightarrow q_j]_{j \in J} : p_1) : i_2 \# (\text{select } [f_k \Rightarrow r_k]_{k \in K} : p_2) : P)} \\
& \xrightarrow{[i_1 \# x!, i_2 \leftarrow i_1 \# x?]} \\
& (S', i_1 \# (q_l \mathrel{++} p_1) : i_2 \# (r_m \mathrel{++} p_2) : P)
\end{aligned}$$

Definition 7 (Scheduling).

$$\begin{aligned}
(Schedule) \quad & \frac{\pi \text{ permutation on } \{1, \dots, n\}}{(S, [i_1 \# p_1, \dots, i_n \# p_n]) \Rightarrow (S, [\pi(i_1) \# p_{\pi(1)}, \dots, \pi(i_n) \# p_{\pi(n)}])} \\
(Closure) \quad & \frac{(S, P) \xRightarrow{T} (S', P') \quad (S', P') \xRightarrow{T'} (S'', P'')}{(S, P) \xRightarrow{T \mathrel{++} T'} (S'', P'')}
\end{aligned}$$

3 Instrumentation and Run-Time Tracing

For each message passing primitive (send/receive) we log two events. In case of send, (1) a *pre* event to indicate the message is about to be sent, and (2) a *post* event to indicate the message has been sent. The treatment is analogous for receive. In our instrumentation, we write $x!$ to denote a single send event and $x?$ to denote a single receive event. These notations are shorthands and can be expressed in terms of the language described so far. We use \equiv to define short-forms and their encodings. We define $x! \equiv [\text{hash}(x), 1]$ and $x? \equiv [\text{hash}(x), 0]$. That is, send is represented by the number 1 and receive by the number 0.

As we support non-deterministic selection, we employ a list of pre events to indicate that one of several events may be chosen. For example, $pre([x!, y?])$ indicates that there is the choice among sending over channel x and receiving over channel y . This is again a shorthand notation where we assume $pre([b_1, \dots, b_n]) \equiv [0, b_1, \dots, b_n]$.

A post event is always singleton as at most one of the possible communications is chosen. As we also trace communication partners, we assume that the sending party transmits its identity, the thread id, to the receiving party. We write $post(i \# x?)$ to denote reception via channel x where the sender has thread id i . In case of a post send event, we simply write $post(x!)$. The above are yet again shorthands where $i \# x? \equiv [\text{hash}(x), 0, i]$ and $post(b) \equiv [1, b]$.

Pre and post events are written in a fresh thread local variable, denoted by x_{tid} where tid refers to the thread's id number. At the start of the thread the variable is initialized by $x_{tid} := []$. Instrumentation ensures that pre and post

events are appropriately logged. As we keep track of communication partners, we must also inject and project messages with additional information (the sender's thread id).

We consider instrumentation of $\text{select } [x \leftarrow 1 \Rightarrow [], y := \leftarrow x \Rightarrow [z \leftarrow y]]$. We assume the above program text is part of a thread with id number 1. We non-deterministically choose between a send and receive operation. In case of receive, the received value is further transmitted. Instrumentation yields the following.

$$\begin{aligned} & [x_1 := x_1 ++ \text{pre}([x!, x?]), \\ & \text{select } [x \leftarrow [\text{tid}, 1] \Rightarrow [x_1 := x_1 ++ \text{post}(x!)], \\ & \quad y' := \leftarrow x \Rightarrow [x_1 := x_1 ++ \text{post}(\text{head}(y')\#x?), y := \text{last}(y'), \\ & \quad z \leftarrow [\text{tid}, y]]] \end{aligned}$$

We first store the pre events, either a read or send via channel x . The send is instrumented by additionally transmitting the senders thread id. The post event for this case simply logs that a send took place. Instrumentation of receive is slightly more involved. As senders supply their thread id, we introduce a fresh variable y' . Via $\text{head}(y')$ we extract the senders thread id to properly record the communication partner in the post event. The actual value transmitted is accessed via $\text{last}(y')$.

Definition 8 (Instrumentation of Programs). *We write $\text{instr}(p) = q$ to denote the instrumentation of program p where q is the result of instrumentation. Function $\text{instr}(\cdot)$ is defined by structural induction on a program. We assume a similar instrumentation function for commands.*

$$\begin{aligned} \text{instr}([]) &= [] \\ \text{instr}(c : p) &= \text{instr}(c) : \text{instr}(p) \\ \text{instr}(y := b) &= [y := b] \\ \text{instr}(y := \text{makeChan}) &= [y := \text{makeChan}] \\ \text{instr}(\text{go } p) &= [\text{go } ([x_{\text{tid}} := [] ++ \text{instr}(p))]] \\ \text{instr}(\text{select } [e_i \Rightarrow p_i]_{i \in \{1, \dots, n\}}) &= [x_{\text{tid}} := x_{\text{tid}} ++ [\text{pre}([\text{retr}(e_1), \dots, \text{retr}(e_n)])], \\ &\quad \text{select } [\text{instr}(e_i \Rightarrow p_i)]_{i \in \{1, \dots, n\}}] \\ \text{instr}(x \leftarrow b \Rightarrow p) &= x \leftarrow [\text{tid}, b] \Rightarrow (x_{\text{tid}} := x_{\text{tid}} ++ [\text{post}(x!)] ++ \text{instr}(p)) \\ \text{instr}(y := \leftarrow x \Rightarrow p) &= y' := \leftarrow x \Rightarrow [x_{\text{tid}} := x_{\text{tid}} ++ [\text{post}(\text{head}(y')\#x?)], \\ &\quad y := \text{last}(y')] ++ \text{instr}(p) \end{aligned}$$

$$\text{retr}(x \leftarrow b) = x! \quad \text{retr}(y := \leftarrow x) = x?$$

Run-time tracing proceeds as follows. We simply run the instrumented program and extract the local traces connected to variables x_{tid} . We assume that thread id numbers are created during program execution and can be enumerated by $1 \dots n$ for some $n > 0$ where thread id number 1 belongs to the main thread.

Definition 9 (Run-Time Tracing). *Let p and q be programs such that $\text{instr}(p) = q$. We consider a specific instrumented program run where $((), [1\# [x_1 := []] ++ q]) \xrightarrow{T}$*

$(S, 1\sharp[] : P)$ for some S , T and P . Then, we refer to T as p 's actual run-time trace. We refer to the list $[1\sharp S(x_1), \dots, n\sharp S(x_n)]$ as the local traces obtained via the instrumentation of p .

Command $x_1 := []$ is added to the instrumented program to initialize the trace of the main thread. Recall that main has thread id number 1. This extra step is necessary because our instrumentation only initializes local traces of threads generated via `go`. The final configuration $(S, 1\sharp[] : P)$ indicates that the main thread has run to full completion. This is a realistic assumption as we assume that programs exhibit no obvious bug during execution. There might still be some pending threads, in case P differs from the empty list.

4 Trace Analysis

We assume that the program has been instrumented and after some program run we obtain a list of local traces. We show that the actual run-time trace can be recovered and we are able to point out alternative behaviors that could have taken place. Alternative behaviors are either due alternative schedules or different choices among communication partners.

We consider the list of local traces $[1\sharp S(x_1), \dots, n\sharp S(x_n)]$. Their shape can be characterized as follows.

Definition 10 (Local Traces).

$$\begin{aligned} U, V &::= [] \mid i\sharp L : U \\ L &::= [] \mid pre(as) : M \\ as &::= [] \mid x! : as \mid x? : as \\ M &::= [] \mid post(x!) : L \mid post(i\sharp x?) : L \end{aligned}$$

We refer to $U = [1\sharp L_1, \dots, n\sharp L_n]$ as a residual list of local traces if for each L_i either $L_i = []$ or $L_i = [pre(\dots)]$.

To recover the communications that took place we check for matching pre and post events recorded in the list of local traces. For this purpose, we introduce a relation $U \xRightarrow{T} V$ to denote that ‘replaying’ of U leads to V where communications T took place. Valid replays are defined via the following rules.

Definition 11 (Replay $U \xRightarrow{T} V$).

$$\begin{aligned}
& \text{(Sync)} \quad \frac{L_1 = \text{pre}([\dots, x!, \dots]) : \text{post}(x!) : L'_1 \quad L_2 = \text{pre}([\dots, x?, \dots]) : \text{post}(i_1 \# x?) : L'_2}{i_1 \# L_1 : i_2 \# L_2 : U \xRightarrow{[i_1 \# x!, i_2 \leftarrow i_1 \# x?]} i_1 \# L'_1 : i_2 \# L'_2 : U} \\
& \text{(Schedule)} \quad \frac{\pi \text{ permutation on } \{1, \dots, n\}}{[i_1 \# L_1, \dots, i_n \# L_n] \xRightarrow{\Pi} [i_{\pi(1)} \# L_{\pi(1)}, \dots, i_{\pi(n)} \# L_{\pi(n)}]} \\
& \text{(Closure)} \quad \frac{U \xRightarrow{T} U' \quad U' \xRightarrow{T'} U''}{U \xRightarrow{T \mathbin{**} T'} U''}
\end{aligned}$$

Rule (Sync) checks for matching communication partners. In each trace, we must find complementary pre events and the post events must match as well. Recall that in the instrumentation the sender transmits its thread id to the receiver. Rule (Schedule) shuffles the local traces as rule (Sync) only considers the two leading local traces. Via rule (Closure) we perform repeated replay steps.

We can state that the actual run-time trace can be obtained via the replay relation $U \xRightarrow{T} V$ but further run-time traces are possible. This is due to alternative schedules.

Proposition 1 (Replay Yields Run-Time Traces). *Let p be a program and q its instrumentation where for a specific program run we observe the actual behavior T and the list $[1 \# L_1, \dots, n \# L_n]$ of local traces. Let $\mathcal{T} = \{T' \mid [1 \# L_1, \dots, n \# L_n] \xRightarrow{T'} 1 \# [] : U \text{ for some residual } U\}$. Then, we find that $T \in \mathcal{T}$ and for each $T' \in \mathcal{T}$ we have that $((), p) \xRightarrow{T'} (S, 1 \# [] : P)$ for some S and P .*

Definition 12 (Alternative Schedules). *We say $[1 \# L_1, \dots, n \# L_n]$ contains alternative schedules iff the cardinality of the set $\{T' \mid [1 \# L_1, \dots, n \# L_n] \xRightarrow{T'} 1 \# [] : U \text{ for some residual } U\}$ is greater than one.*

We can also check if even further run-time traces might have been possible by testing for alternative communications.

Definition 13 (Alternative Communications). *We say $[1 \# L_1, \dots, n \# L_n]$ contains alternative matches iff for some i, j, x, L, L' we have that (1) $L_i = \text{pre}([\dots, x!, \dots]) : L$, (2) $L_j = \text{pre}([\dots, x?, \dots]) : L'$, and (3) if $L = \text{post}(x!) : L''$ for some L'' then $L' \neq \text{post}(j \# x?) : L'''$ for any L''' .*

We say $U = [1 \# L_1, \dots, n \# L_n]$ contains alternative communications iff U contains alternative matches or there exists T and V such that $U \xRightarrow{T} V$ and V contains alternative matches.

The alternative match condition states that a sender could synchronize with a receiver (see (1) and (2)) but this synchronization did not take place (see (3)). For an alternative match to result in an alternative communication, the match must be along a possible run-time trace.

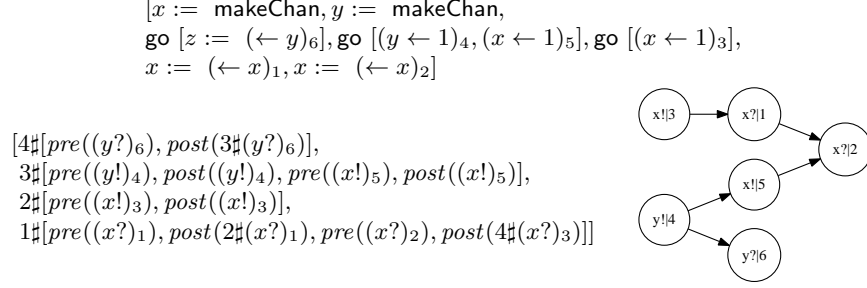


Fig. 1: Dependency Graph among Events

4.1 Dependency Graph for Efficient Trace Analysis

Instead of replaying traces to check for alternative schedules and communications, we build a dependency graph where the graph captures the partial order among events. It is much more efficient to carry out the analysis on the graph than replaying traces. Figure 1 shows a simple example.

We find a program that makes use of two channels and four threads. For reference, send/receive events are annotated (as subscript) with unique numbers. We omit the details of instrumentation and assume that for a specific program run we find the list of given traces on the left. Pre events consist of singleton lists as there is no `select`. Hence, we write $pre((y?)_6)$ as a shorthand for $pre([(y?)_6])$. Replay of the trace shows that the following locations synchronize with each other: (4, 6), (3, 1) and (5, 2). This information as well as the order among events can be captured by a dependency graph. Nodes are obtained by a linear scan through the list of traces. To derive edges, we require another scan for each element in a trace as we need to find pre/post pairs belonging to matching synchronizations. This results overall in $O(m * m)$ for the construction of the graph where m is the number of elements found in each trace. To avoid special treatment of dangling pre events (with not subsequent post event), we assume that some dummy post events are added to the trace.

Definition 14 (Construction of Dependency Graph). *Each node corresponds to a send or a receive operation in the program text. Edges are constructed by observing events recorded in the list of traces. We draw a (directed) edge among nodes if either*

- *the pre and post events of one node precede the pre and post events of another node in the trace, or*
- *the pre and post events belonging to both nodes can be synchronized. See rule (Sync) in Definition 11. We assume that the edge starts from the node with the send operation.*

Applied to our example, this results in the graph on the right. See Figure 1. For example, $x!|3$ denotes a send communication over channel x at program

location 3. As send precedes receive we find an edge from $x!|3$ to $x?|1$. In general, there may be several initial nodes. By construction, each node has at most one outgoing edge but may have multiple incoming edges.

The trace analysis can be carried out directly on the dependency graph. To check if one event happens-before another event we seek for a path from one event to the other. This can be done via a depth-first search and takes time $O(v + e)$ where v is the number of nodes and e the number of edges. Two events are concurrent if neither happens-before the other. To check for alternative communications, we check for matching nodes that are concurrent to each other. By matching we mean that one of the nodes is a send and the other is a receive over the same channel. For our example, we find that $x!|5$ and $x?|1$ represents an alternative communication as both nodes are matching and concurrent to each other.

To derive (all) alternative schedules, we perform a backward traversal of the graph. Backward in the sense that we traverse the graph by moving from children to parent node. We start with some final node (no outgoing edge). Each node visited is marked. We proceed to the parent if all children are marked. Thus, we guarantee that the happens-before relation is respected. For our example, suppose we visit first $y?6$. We cannot visit its parent $y!4$ until we have visited $x?2$ and $x!5$. Via a (backward) breadth-first search we can ‘accumulate’ all schedules.

5 Comparison to Vector Clock Method

Via a simple adaptation of the Replay Definition 11 we can attach vector clocks to each send and receive event. Hence, our tracing method strictly subsumes the vector clock method as we are also able to trace events that could not commit.

Definition 15 (Vector Clock).

$$cs ::= [] \mid n : cs$$

For convenience, we represent a vector clock as a list of clocks where the first position belongs to thread 1 etc. We write $cs[i]$ to retrieve the i -th component in cs . We write $\text{inc}(i, cs)$ to denote the vector clock obtained from cs where all elements are the same but at index i the element is incremented by one. We write $\text{max}(cs_1, cs_2)$ to denote the vector clock where we per-index take the greater element. We write i^{cs} to denote thread i with vector clock cs . We write $i \# x!^{cs}$ to denote a send over channel x in thread i with vector clock cs . We write $i \leftarrow j \# x?^{cs}$ to denote a receive over channel x in thread i from thread j with vector clock cs .

Definition 16 (From Trace Replay to Vector Clocks).

$$\begin{array}{c} L_1 = \text{pre}([\dots, x!, \dots]) : \text{post}(x!) : L'_1 \\ L_2 = \text{pre}([\dots, x?, \dots]) : \text{post}(i_1 \# x?) : L'_2 \\ \text{(Sync)} \quad cs = \text{max}(\text{inc}(i_1, cs_1), \text{inc}(i_2, cs_2)) \\ \hline i_1^{cs_1} \# L_1 : i_2^{cs_2} \# L_2 : U \xrightarrow{[i_1 \# x!^{cs}, i_2 \leftarrow i_1 \# x?^{cs}]} i_1^{cs} \# L'_1 : i_2^{cs} \# L'_2 : U \end{array}$$

Like the construction of the dependency graph, the (re)construction of vector clocks takes time $O(m * m)$ where m is the number of elements found in each trace.

To check for an alternative communication, the vector clock method seeks for matching events. This incurs the same (quadratic in the size of the trace) cost as for our method. However, the check that these two events are concurrent to each other can be performed more efficiently via vector clocks. Comparison of vector clocks takes time $O(n)$ where n is the number of threads. Recall that our graph-based method requires time $O(v + e)$ where v is the number of nodes and e the number of edges. The number n is smaller than $v + e$.

However, our dependency graph representation is more efficient in case of exploring alternative schedules. In case of the vector clock method, we need to continuously compare vector clocks whereas we only require a (backward) traversal of the graph. We believe that the dependency graph has further advantages in case of user interaction and visualization as it is more intuitive to navigate through the graph. This is something we intend to investigate in future work.

6 Implementation

We have fully integrated the approach laid out in the earlier sections into the Go programming language and have built a prototype tool. We give an overview of our implementation which can be found here [5]. A detailed treatment of all of Go’s message-passing features can be found in the extended version of this paper.

6.1 Library-Based Instrumentation and Tracing

We use a pre-processor to carry out the instrumentation as described in Section 3. In our implementation, each thread maintains an entry in a lock-free hashmap where each entry represents a thread (trace). The hashmap is written to file either at the end of the program or when a deadlock occurs. We currently do not deal with the case that the program crashes as we focus on the detection of potential bugs in programs that do not show any abnormal behavior.

6.2 Measurement of Run-Time Overhead Library-Based Tracing

We measure the run-time overhead of our method against the vector clock method. Both methods are implemented as libraries assuming no access to the Go run-time system. For experimentation we use three programs where each program exercises some of the factors that have an impact on tracing. For example, dynamic versus static number of threads and channels. Low versus high amount of communication among threads.

The **Add-Pipe (AP)** example uses n threads where the first $n - 1$ threads receive on an input channel, add one to the received value and then send the

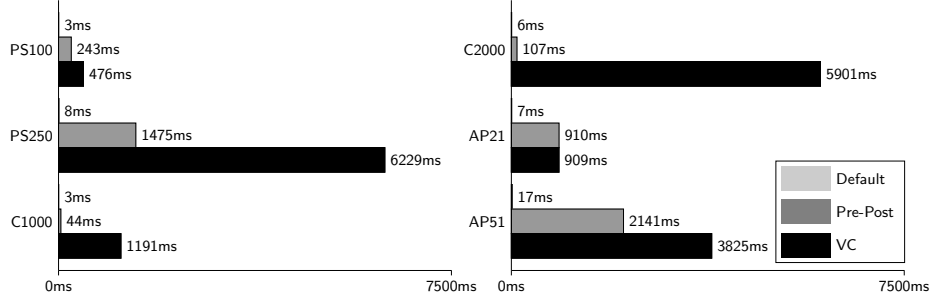


Fig. 2: Performance overhead using Pre/Post vs Vector clocks(VC) in ms.

new value on their output channel to the next thread. The first thread sends the initial value and receives the result from the last thread.

In the **Primesieve (PS)** example, the communication among threads is similar to the **Add-Pipe** example. The difference is that threads and channels are dynamically generated to calculate the first n prime numbers. For each found prime number a ‘filter’ thread is created. Each thread has an input channel to receive new possible prime numbers v and an output channel to report each number for which $v \bmod prime \neq 0$ where $prime$ is the prime number associated with this filter thread. The filter threads are run in a chain where the first thread stores the prime number 2.

The **Collector (C)** example creates n threads that produce a number which is then sent to the main thread for collection. This example has much fewer communications compared to the other examples but uses a high number of threads.

Figure 2 summarizes our results. Results are carried out on some commodity hardware (Intel i7-6600U with 12 GB RAM, a SSD and Go 1.8.3 running on Windows 10 was used for the tests). Our results show that a library-based implementation of the vector clock method does not scale well for examples with a dynamic number of threads and/or a high amount communication among threads. See examples **Primesieve** and **Add-Pipe**. None of the vector clock optimizations [3] apply here because of the dynamic number of threads and channels. Our method performs much better. This is no surprise as we require less (tracing) data and no extra communication links. We believe that the overhead can still be further reduced as access to the thread id in Go is currently rather cumbersome and expensive.

7 Conclusion

One of the challenges of run-time verification in the concurrent setting is to establish a partial order among recorded events. Thus, we can identify potential bugs due to bad schedules that are possible but did not take place in some specific program run. Vector clocks are the predominant method to achieve this task. For

example, see work by Vo [11] in the MPI setting and work by Tasharofi [10] in the actor setting. There are several works that employ vector clocks in the shared memory setting. For example, see Pozniansky’s and Schuster’s work [9] on data race detection. Some follow-up work by Flanagan and Freund [2] employs some optimizations to reduce the tracing overhead by recording only a single clock instead of the entire vector. We leave to future work to investigate whether such optimizations are applicable in the message-passing setting and how they compare to existing optimizations such as [3].

We have introduced a novel tracing method that has much less overhead compared to the vector clock method. Our method can deal with all of Go’s message-passing language features and can be implemented efficiently as a library. We have built a prototype that can automatically identify alternative schedules and communications. In future work we plan to conduct some case studies and integrate heuristics for specific scenarios, e.g. reporting a send operation on a closed channel etc.

Acknowledgments

We thank some HVC’17 reviewers for their constructive feedback on an earlier version of this paper.

References

1. C. J. Fidge. Timestamps in message-passing systems that preserve the partial ordering. 10(1):56–66, 1987.
2. C. Flanagan and S. N. Freund. Fasttrack: Efficient and precise dynamic race detection. In *Proc. of PLDI ’09*, pages 121–133. ACM, 2009.
3. V. K. Garg, C. Skawratananond, and N. Mittal. Timestamping messages and events in a distributed system using synchronous communication. *Distributed Computing*, 19(5-6):387–402, 2007.
4. The Go programming language. <https://golang.org/>.
5. Trace-based run-time analysis of message-passing Go programs. <https://github.com/KaiSta/gopherlyzer-GoScout>.
6. C. A. R. Hoare. Communicating sequential processes. *Commun. ACM*, 21(8):666–677, Aug. 1978.
7. L. Lamport. Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7):558–565, 1978.
8. F. Mattern. Virtual time and global states of distributed systems. In *Parallel and Distributed Algorithms*, pages 215–226. North-Holland, 1989.
9. E. Pozniansky and A. Schuster. Multirace: efficient on-the-fly data race detection in multithreaded C++ programs. *Concurrency and Computation: Practice and Experience*, 19(3):327–340, 2007.
10. S. Tasharofi. *Efficient testing of actor programs with non-deterministic behaviors*. PhD thesis, University of Illinois at Urbana-Champaign, 2013.
11. A. Vo. *Scalable Formal Dynamic Verification of Mpi Programs Through Distributed Causality Tracking*. PhD thesis, University of Utah, 2011. AAI3454168.


```

func A(x chan int) {
    x <- 1          // A1
}

func bufferedChan() {
    x := make(chan int, 1)
    go A(x)
    x <- 1          // A2
    <-x
}

func closedChan() {
    x := make(chan int)
    go A(x)
    go B(x)
    close(x)
}

func B(x chan int) {
    <-x
}

func selDefault() {
    x := make(chan int)
    go A(x)

    select {
    case <-x:          // A3
        fmt.Println("received from x")
    default:
        fmt.Println("default")
    }
}

```

Fig. 3: Further Go Features

A Further Go Message-Passing Features

A.1 Overview

Besides selective synchronous message-passing, Go supports some further message passing features that can be easily dealt with by our approach and are fully supported by our implementation. Figure 3 shows such examples where we put the program text in two columns.

Buffered Channels Go also supports buffered channels where send is asynchronous assuming sufficient buffer space exists. See function `buffered` in Figure 3. Depending on the program run, our analysis reports that either A1 or A2 are alternative matches for the receive operation.

In terms of the instrumentation and tracing, we treat each asynchronous send as if the send is executed in its own thread. This may lead to some slight inaccuracies. Consider the following variant.

```

func buffered2() {
    x := make(chan int, 1)
    x <- 1          // B1
    go A(x)         // B2
    <-x              // B3
}

```

Our analysis reports that B2 and B3 form an alternative match. However, in the Go semantics, buffered messages are queued. Hence, for *every* program run the only possibility is that B1 synchronizes with B3. B3 never takes place! As our

main objective is bug finding, we argue that this loss of accuracy is justifiable. How to eliminate such false positives is subject of future work.

Select with default/timeout Another feature in Go is to include a default/timeout case to `select`. See `selDefault` in Figure 3. The purpose is to avoid (indefinite) blocking if none of the other cases are available. For the user it is useful to find out if other alternatives are available in case the default case is selected. The default case applies for most program runs. Our analysis reports that A1 and A3 are an alternative match.

To deal with default/timeout we introduce a new post event $post(select)$. To carry out the analysis in terms of the dependency graph, each subtrace $\dots, pre([\dots, select, \dots]), post(select), \dots$ creates a new node. Construction of edges remains unchanged.

Closing of Channels Another feature in Go is the ability to close a channel. See `closedChan` in Figure 3. Once a channel is closed, each send on a closed channel leads to failure (the program crashes). On the other hand, each receive on a closed channel is always successful, as we receive a dummy value. A run of is successful if the close operation of the main thread happens after the send in thread A. As the close and send operations happen concurrently, our analysis reports that the send A1 may take place after close.

For instrumentation/tracing, we introduce event $close(x)$. It is easy to identify a receive on a closed channel, as we receive a dummy thread id. So, for each subtrace $[\dots, pre([\dots, x?, \dots]), post(i\#x?), \dots]$ where i is a dummy value we draw an edge from $close(x)$ to $x?$.

Here are the details of how to include buffered channels, select and closing of channels.

A.2 Buffered Channels

Consider the following Go program.

```
x := make(chan, 2)
x <- 1      // E1
x <- 1      // E2
<- x       // E3
<- x       // E4
```

We create a buffer of size 2. The two send operations will then be carried out asynchronously and the subsequent receive operations will pick up the buffered values. We need to take special care of buffered send operations. If we would treat them like synchronous send operations, their respective pre and post events would be recorded in the same trace as the pre and post events of the receive operations. This would have the consequence that our trace analysis does not find out that events E1 and E2 happen before E3 and E4.

Our solution to this issue is to treat each send operation on a buffered channel as if the send operation is carried out in its own thread. Thus, our trace analysis

is able to detect that E1 and E2 take place before E3 and E4. This is achieved by marking each send on a buffered channel in the instrumentation. After tracing, pre and post events will then be moved to their own trace. From the viewpoint of our trace analysis, a buffered channel then appears as having infinite buffer space. Of course, when running the program a send operation may still block if all buffer space is occupied.

Here are the details of the necessary adjustments to our method. During instrumentation/tracing, we simply record if a buffered send operation took place. The only affected case in the instrumentation of commands (Definition 8) is $x \leftarrow b \Rightarrow p$. We assume a predicate $\text{isBuffered}(\cdot)$ to check if a channel is buffered or not. In terms of the actual implementation this is straightforward to implement. We write $\text{postB}(x, n)$ to indicate a buffered send operation via x where n is a fresh thread id. We create fresh thread id numbers via tidB .

Definition 17 (Instrumentation of Buffered Channels). *Let x be a buffered channel.*

$$\begin{aligned} \text{instr}(x \leftarrow b \Rightarrow p) \\ &| \text{isBuffered}(x) = x \leftarrow [n, b] \Rightarrow (x_{\text{tid}} := x_{\text{tid}} \mathrel{++} [\text{postB}(x!n,)]) \mathrel{++} \text{instr}(p) \\ &\quad \text{where } n = \text{tidB} \\ &| \text{otherwise} = x \leftarrow [\text{tid}, b] \Rightarrow (x_{\text{tid}} := x_{\text{tid}} \mathrel{++} [\text{post}(x!)]) \mathrel{++} \text{instr}(p) \end{aligned}$$

The treatment of buffered channels has no overhead on the instrumentation and tracing. However, we require a post-processing phase where marked events will be then moved to their own trace. This can be achieved via a linear scan through each trace. Hence, requires time complexity $O(k)$ where k is the overall size of all (initially recorded) traces. For the sake of completeness, we give below a declarative description of post-processing in terms of relation $U \Rightarrow V$.

Definition 18 (Post-Processing for Buffered Channels $U \Rightarrow V$).

$$\begin{aligned} (\text{MovePostB}) \quad & \frac{L = \text{pre}(as) : \text{postB}(x!, n) : L'}{i\#L : U \Rightarrow i\#L' : n\#[\text{pre}(as), \text{postB}(x!, n)] : U} \\ (\text{Shift}) \quad & \frac{\begin{array}{c} L = \text{pre}(as) : \text{post}(a) : L' \\ (a = x! \vee a = j\#x?) \\ i\#L' : U \Rightarrow i\#L'' : U' \end{array}}{i\#L : U \Rightarrow i\#\text{pre}(as) : \text{post}(a) : L'' : U'} \\ (\text{Schedule}) \quad & \frac{\pi \text{ permutation on } \{1, \dots, n\}}{[i_1\#L_1, \dots, i_n\#L_n] \Rightarrow [i_{\pi(1)}\#L_{\pi(1)}, \dots, i_{\pi(n)}\#L_{\pi(n)}]} \\ (\text{Closure}) \quad & \frac{U \Rightarrow U' \quad U' \Rightarrow U''}{U \Rightarrow U''} \end{aligned}$$

Subsequent analysis steps will be carried out on the list of traces obtained via post-processing.

There is some space for improvement. Consider the following program text.

```
func A(x chan int) {
  x <- 1          // A1
}
func buffered2() {
  x := make(chan int, 1)
  x <- 1          // B1
  go A(x)         // B2
  <-x             // B3
}
```

Our analysis (for some program run) reports that B2 and B3 is an alternative match. However, in the Go semantics, buffered messages are queued. Hence, for *every* program run the only possibility is that B1 synchronizes with B3. B3 never takes place. As our main objective is bug finding, we can live with this inaccuracy. We will investigate in future work how to eliminate this false positive.

B Select with default/timeout

In terms of the instrumentation/tracing, we introduce a new special post event $post(select)$. For the trace analysis (Definition 11), we require a new rule.

$$(\text{Default/Timeout}) \ i\sharp pre([\dots]) : post(select) : L : U \xrightarrow{i\sharp select} i\sharp L : U$$

This guarantees that in case default or timeout is chosen, select acts as if asynchronous.

The dependency graph construction easily takes care of this new feature. For each default/timeout case we introduce a node. Construction of edges remains unchanged.

C Closing of Channels

For instrumentation/tracing of the $close(x)$ operation on channel x , we introduce a special pre and post event. Our trace analysis keeps track of closed channels. As a receive on a closed channel yields some dummy values, it is easy to distinguish this case from the regular (Sync). Here are the necessary adjustments to our replay relation from Definition 11.

$$C ::= [] \mid i\sharp close(x) : C$$

$$(\text{Close}) \ (i\sharp\text{pre}(\text{close}(x)) : \text{post}(\text{close}(x)) : L : U \mid C) \Downarrow (i\sharp L : U \mid i\sharp\text{close}(x) : C)$$

$$(\text{RcvClosed}) \ \frac{Q = j\sharp\text{close}(x) : Q'}{(i\sharp\text{pre}([\dots, x?, \dots]) : \text{post}(j'\sharp x?) : L : U \mid Q) \xrightarrow{[j\sharp\text{close}(x), i\leftarrow j\sharp x?]} (i\sharp L : U \mid Q)}$$

For the construction of the dependency graph, we create a node for each close statement. For each receive on a closed channel x at program location l , we draw an edge from $\text{close}(x)$ to $x?|l$.

D Codes used for the Experimental results

Add-Pipe

```
func add1(in chan int) chan int {
    out := make(chan int)
    go func() {
        for {
            n := <-in
            out <- n + 1
        }
    }()
    return out
}

func main() {
    in := make(chan int)
    c1 := add1(in)
    for i := 0; i < 19; i++ {
        c1 = add1(c1)
    }
    for n := 1; n < 1000; n++ {
        in <- n
        <-c1
    }
}
```

Primesieve

```
func generate(ch chan int) {
    for i := 2; ; i++ {
        ch <- i
    }
}
```

```

    }
}

func filter(in chan int, out chan int, prime int) {
    for {
        tmp := <-in
        if tmp%prime != 0 {
            out <- tmp
        }
    }
}

func main() {
    ch := make(chan int)
    go generate(ch)
    for i := 0; i < 100; i++ {
        prime := <-ch
        ch1 := make(chan int)
        go filter(ch, ch1, prime)
        ch = ch1
    }
}

```

Collector

```

func collect(x chan int, v int) {
    x <- v
}

func main() {
    x := make(chan int)
    for i := 0; i < 1000; i++ {
        go collect(x, i)
    }

    for i := 0; i < 1000; i++ {
        <-x
    }
}

```

