Ruby master - Bug #13299

backport r57469, r57472, r57508 (garbage reduction for IO#write/syswrite)

03/10/2017 07:46 PM - normalperson (Eric Wong)

Status:	Closed		
Priority:	Normal		
Assignee:			
Target version:			
ruby -v:		**Backport:**	2.2: REQUIRED, 2.3: REQUIRED, 2.4: DONE

Description

This is a performance regression introduced in Ruby 2.0.0 and maybe some folks would like it fixed in a stable release before December 25, 2017

Thank you.

Related issues:

Related to Ruby master - Bug #13085: io.c io_fwrite creates garbage	**Closed**

Associated revisions

Revision baf330be - 03/12/2017 06:15 PM - naruse (Yui NARUSE)

merge revision(s) 57469,57471,57472,57503,57508: [Backport #13299]

```
    io.c: recycle garbage on write

    * string.c (STR_IS_SHARED_M): new flag to mark shared mulitple times
      (STR_SET_SHARED): set STR_IS_SHARED_M
      (rb_str_tmp_frozen_acquire, rb_str_tmp_frozen_release): new functions
      (str_new_frozen): set/unset STR_IS_SHARED_M as appropriate
    * internal.h: declare new functions
    * io.c (fwrite_arg, fwrite_do, fwrite_end): new
      (io_fwrite): use new functions

    Introduce rb_str_tmp_frozen_acquire and rb_str_tmp_frozen_release
    to manage a hidden, frozen string.  Reuse one bit of the embed
    length for shared strings as STR_IS_SHARED_M to indicate a string
    has been shared multiple times.  In the common case, the string
    is only shared once so the object slot can be reclaimed immediately.

    minimum results in each 3 measurements. (time and size)

    Execution time (sec)
    name                             trunk   built
    io_copy_stream_write             0.682   0.254
    io_copy_stream_write_socket      1.225   0.751

    Speedup ratio: compare with the result of `trunk' (greater is better)
    name    built
    io_copy_stream_write             2.680
    io_copy_stream_write_socket      1.630

    Memory usage (last size) (B)
    name                             trunk           built
    io_copy_stream_write             95436800.000    6512640.000
    io_copy_stream_write_socket      117628928.000   7127040.000

    Memory consuming ratio (size) with the result of `trunk' (greater is better)
    name    built
    io_copy_stream_write             14.654
    io_copy_stream_write_socket      16.505
    string.c (rb_str_tmp_frozen_release): release embedded strings

    Handle the embedded case first, since we may have an embedded
    duplicate and non-embedded original string.
    * string.c (rb_str_tmp_frozen_release): handled embedded strings
    * test/ruby/test_io.rb (test_write_no_garbage): new test
      [ruby-core:78898] [Bug #13085]
    io.c (rb_io_syswrite): avoid leaving garbage after write

    As with IO#write, IO#syswrite also generates garbage which can
    be harmful in hand-coded read-write loops.

    * io.c (swrite_arg, swrite_do, swrite_end): new
      (rb_io_syswrite): use new functions to cleanup garbage
      [ruby-core:78898] [Bug #13085]
    Add class name to assert messages
    io.c: remove rb_ensure usage for rb_str_tmp_frozen_* calls

    Using rb_ensure pessimizes the common case and makes the code
    more difficult to read and follow.  If we hit an exceptions
    during write, just let the GC handle cleanup as the exception
    is already bad for garbage.

    * io.c (io_fwrite): call rb_str_tmp_frozen{acquire,release} directly
      (rb_io_syswrite): ditto
      (fwrite_do, fwrite_end, swrite_do, swrite_end): remove
```

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_4@57941 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 57941 - 03/12/2017 06:15 PM - naruse (Yui NARUSE)

merge revision(s) 57469,57471,57472,57503,57508: [Backport #13299]

```
io.c: recycle garbage on write

* string.c (STR_IS_SHARED_M): new flag to mark shared mulitple times
  (STR_SET_SHARED): set STR_IS_SHARED_M
  (rb_str_tmp_frozen_acquire, rb_str_tmp_frozen_release): new functions
  (str_new_frozen): set/unset STR_IS_SHARED_M as appropriate
* internal.h: declare new functions
* io.c (fwrite_arg, fwrite_do, fwrite_end): new
  (io_fwrite): use new functions

Introduce rb_str_tmp_frozen_acquire and rb_str_tmp_frozen_release
to manage a hidden, frozen string.  Reuse one bit of the embed
length for shared strings as STR_IS_SHARED_M to indicate a string
has been shared multiple times.  In the common case, the string
is only shared once so the object slot can be reclaimed immediately.

minimum results in each 3 measurements. (time and size)

Execution time (sec)
name                            trunk   built
io_copy_stream_write            0.682   0.254
io_copy_stream_write_socket     1.225   0.751

Speedup ratio: compare with the result of `trunk' (greater is better)
name    built
io_copy_stream_write            2.680
io_copy_stream_write_socket     1.630

Memory usage (last size) (B)
name                            trunk           built
io_copy_stream_write            95436800.000    6512640.000
io_copy_stream_write_socket     117628928.000   7127040.000

Memory consuming ratio (size) with the result of `trunk' (greater is better)
name    built
io_copy_stream_write            14.654
io_copy_stream_write_socket     16.505
string.c (rb_str_tmp_frozen_release): release embedded strings

Handle the embedded case first, since we may have an embedded
duplicate and non-embedded original string.

* string.c (rb_str_tmp_frozen_release): handled embedded strings
* test/ruby/test_io.rb (test_write_no_garbage): new test
  [ruby-core:78898] [Bug #13085]
io.c (rb_io_syswrite): avoid leaving garbage after write
```

```
As with IO#write, IO#syswrite also generates garbage which can
be harmful in hand-coded read-write loops.

* io.c (swrite_arg, swrite_do, swrite_end): new
  (rb_io_syswrite): use new functions to cleanup garbage
  [ruby-core:78898] [Bug #13085]
Add class name to assert messages
io.c: remove rb_ensure usage for rb_str_tmp_frozen_* calls

Using rb_ensure pessimizes the common case and makes the code
more difficult to read and follow.  If we hit an exceptions
during write, just let the GC handle cleanup as the exception
is already bad for garbage.

* io.c (io_fwrite): call rb_str_tmp_frozen{acquire,release} directly
  (rb_io_syswrite): ditto
  (fwrite_do, fwrite_end, swrite_do, swrite_end): remove
```

History

#1 - 03/10/2017 07:54 PM - normalperson (Eric Wong)

- Status changed from Open to Closed

Closing to trigger backport request

#2 - 03/12/2017 06:06 PM - naruse (Yui NARUSE)

- Related to Bug #13085: io.c io_fwrite creates garbage added

#3 - 03/12/2017 06:15 PM - naruse (Yui NARUSE)

- Backport changed from 2.2: REQUIRED, 2.3: REQUIRED, 2.4: REQUIRED to 2.2: REQUIRED, 2.3: REQUIRED, 2.4: DONE

ruby_2_4 r57941 merged revision(s) 57469,57471,57472,57503,57508.