

```

*****
22310 Thu Oct 30 22:02:12 2014
new/usr/src/cmd/sgs/libld/common/unwind.c
5270 ld(1) cannot handle CIE version 3 in .eh_frame
*****
1 /*
2  * CDDL HEADER START
3  *
4  * The contents of this file are subject to the terms of the
5  * Common Development and Distribution License (the "License").
6  * You may not use this file except in compliance with the License.
7  *
8  * You can obtain a copy of the license at usr/src/OPENSOLARIS.LICENSE
9  * or http://www.opensolaris.org/os/licensing.
10 * See the License for the specific language governing permissions
11 * and limitations under the License.
12 *
13 * When distributing Covered Code, include this CDDL HEADER in each
14 * file and include the License file at usr/src/OPENSOLARIS.LICENSE.
15 * If applicable, add the following below this CDDL HEADER, with the
16 * fields enclosed by brackets "[]" replaced with your own identifying
17 * information: Portions Copyright [yyyy] [name of copyright owner]
18 *
19 * CDDL HEADER END
20 */

22 /*
23  * Copyright (c) 2005, 2010, Oracle and/or its affiliates. All rights reserved.
24  * Copyright 2014 Nexenta Systems, Inc.
25  */

27 #include <string.h>
28 #include <stdio.h>
29 #include <sys/types.h>
30 #include <sgs.h>
31 #include <debug.h>
32 #include <libld.h>
33 #include <dwarf.h>
34 #include <stdlib.h>

36 /*
37  * A EH_FRAME_HDR consists of the following:
38  *
39  *      Encoding      Field
40  *      -----
41  *      unsigned byte  version
42  *      unsigned byte  eh_frame_ptr_enc
43  *      unsigned byte  fde_count_enc
44  *      unsigned byte  table_enc
45  *      encoded        eh_frame_ptr
46  *      encoded        fde_count
47  *      [ binary search table ]
48  *
49  * The binary search table entries each consists of:
50  *
51  *      encoded        initial_func_loc
52  *      encoded        FDE_address
53  *
54  * The entries in the binary search table are sorted
55  * in a increasing order by the initial location.
56  *
57  *
58  * version
59  *
60  * Version of the .eh_frame_hdr format. This value shall be 1.
61  */

```

```

62 * eh_frame_ptr_enc
63 *
64 * The encoding format of the eh_frame_ptr field. For shared
65 * libraries the encoding must be
66 * DW_EH_PE_sdata4|DW_EH_PE_pcrel or
67 * DW_EH_PE_sdata4|DW_EH_PE_datarel.
68 *
69 *
70 * fde_count_enc
71 *
72 * The encoding format of the fde_count field. A value of
73 * DW_EH_PE_omit indicates the binary search table is not
74 * present.
75 *
76 * table_enc
77 *
78 * The encoding format of the entries in the binary search
79 * table. A value of DW_EH_PE_omit indicates the binary search
80 * table is not present. For shared libraries the encoding
81 * must be DW_EH_PE_sdata4|DW_EH_PE_pcrel or
82 * DW_EH_PE_sdata4|DW_EH_PE_datarel.
83 *
84 *
85 * eh_frame_ptr
86 *
87 * The encoded value of the pointer to the start of the
88 * .eh_frame section.
89 *
90 * fde_count
91 *
92 * The encoded value of the count of entries in the binary
93 * search table.
94 *
95 * binary search table
96 *
97 * A binary search table containing fde_count entries. Each
98 * entry of the table consist of two encoded values, the
99 * initial location of the function to which an FDE applies,
100 * and the address of the FDE. The entries are sorted in an
101 * increasing order by the initial location value.
102 *
103 */

106 /*
107  * EH_FRAME sections
108  * =====
109  *
110  * The call frame information needed for unwinding the stack is output in
111  * an ELF section(s) of type SHT_AMD64_UNWIND (amd64) or SHT_PROGBITS (other).
112  * In the simplest case there will be one such section per object file and it
113  * will be named ".eh_frame". An .eh_frame section consists of one or more
114  * subsections. Each subsection contains a CIE (Common Information Entry)
115  * followed by varying number of FDEs (Frame Descriptor Entry). A FDE
116  * corresponds to an explicit or compiler generated function in a
117  * compilation unit, all FDEs can access the CIE that begins their
118  * subsection for data.
119  *
120  * If an object file contains C++ template instantiations, there shall be
121  * a separate CIE immediately preceding each FDE corresponding to an
122  * instantiation.
123  *
124  * Using the preferred encoding specified below, the .eh_frame section can
125  * be entirely resolved at link time and thus can become part of the
126  * text segment.
127  */

```

```

128 * .eh_frame Section Layout
129 * -----
130 *
131 * EH_PE encoding below refers to the pointer encoding as specified in the
132 * enhanced LSB Chapter 7 for Eh_Frame_Hdr.
133 *
134 * Common Information Entry (CIE)
135 * -----
136 * CIE has the following format:
137 *
138 *           Length
139 *           in
140 *           Field      Byte      Description
141 *           ----      -
142 * 1. Length          4          Length of CIE (not including
143 *                    this 4-byte field).
144 *
145 * 2. CIE id          4          Value Zero (0) for .eh_frame
146 *                    (used to distinguish CIEs and
147 *                    FDEs when scanning the section)
148 *
149 * 3. Version          1          Value One (1)
150 *
151 * 4. CIE Augmentation string    Null-terminated string with legal
152 *                    values being "" or 'z' optionally
153 *                    followed by single occurrences of
154 *                    'P', 'L', or 'R' in any order.
155 *                    The presence of character(s) in the
156 *                    string dictates the content of
157 *                    field 8, the Augmentation Section.
158 *                    Each character has one or two
159 *                    associated operands in the AS.
160 *                    Operand order depends on
161 *                    position in the string ('z' must
162 *                    be first).
163 *
164 * 5. Code Align Factor  uleb128  To be multiplied with the
165 *                    "Advance Location" instructions in
166 *                    the Call Frame Instructions
167 *
168 * 6. Data Align Factor  sleb128  To be multiplied with all offset
169 *                    in the Call Frame Instructions
170 *
171 * 7. Ret Address Reg    1          A "virtual" register representation
172 *                    of the return address. In Dwarf V2,
173 *                    this is a byte, otherwise it is
174 *                    uleb128. It is a byte in gcc 3.3.x
175 *
176 * 8. Optional CIE      varying   Present if Augmentation String in
177 *                    Augmentation Section
178 *                    field 4 is not 0.
179 *
180 *   z:
181 *   size                uleb128  Length of the remainder of the
182 *                    Augmentation Section
183 *
184 *   P:
185 *   personality_enc     1          Encoding specifier - preferred
186 *                    value is a pc-relative, signed
187 *                    4-byte
188 *
189 *   personality routine (encoded) Encoded pointer to personality
190 *                    routine (actually to the PLT
191 *                    entry for the personality
192 *                    routine)
193 *
194 *   R:

```

```

194 *   code_enc           1          Non-default encoding for the
195 *                    code-pointers (FDE members
196 *                    "initial_location" and "address_range"
197 *                    and the operand for DW_CFA_set_loc)
198 *                    - preferred value is pc-relative,
199 *                    signed 4-byte.
200 *
201 *   L:
202 *   lsda_enc           1          FDE augmentation bodies may contain
203 *                    LSDA pointers. If so they are
204 *                    encoded as specified here -
205 *                    preferred value is pc-relative,
206 *                    signed 4-byte possibly indirect
207 *                    thru a GOT entry.
208 *
209 * 9. Optional Call Frame varying
210 *   Instructions
211 *
212 * The size of the optional call frame instruction area must be computed
213 * based on the overall size and the offset reached while scanning the
214 * preceding fields of the CIE.
215 *
216 *
217 * Frame Descriptor Entry (FDE)
218 * -----
219 * FDE has the following format:
220 *
221 *           Length
222 *           in
223 *           Field      Byte      Description
224 *           ----      -
225 * 1. Length          4          Length of remainder of this FDE
226 *
227 * 2. CIE Pointer          4          Distance from this field to the
228 *                    nearest preceding CIE
229 *                    (uthe value is subtracted from the
230 *                    current address). This value
231 *                    can never be zero and thus can
232 *                    be used to distinguish CIE's and
233 *                    FDE's when scanning the
234 *                    .eh_frame section
235 *
236 * 3. Initial Location    varying   Reference to the function code
237 *                    corresponding to this FDE.
238 *                    If 'R' is missing from the CIE
239 *                    Augmentation String, the field is an
240 *                    8-byte absolute pointer. Otherwise,
241 *                    the corresponding EH_PE encoding in the
242 *                    CIE Augmentation Section is used to
243 *                    interpret the reference.
244 *
245 * 4. Address Range        varying   Size of the function code corresponding
246 *                    to this FDE.
247 *                    If 'R' is missing from the CIE
248 *                    Augmentation String, the field is an
249 *                    8-byte unsigned number. Otherwise,
250 *                    the size is determined by the
251 *                    corresponding EH_PE encoding in the
252 *                    CIE Augmentation Section (the
253 *                    value is always absolute).
254 *
255 * 5. Optional FDE      varying   present if CIE augmentation
256 *                    Augmentation Section string is non-empty.
257 *
258 *
259 *   'z':

```

```

260 *      length          uleb128      length of the remainder of the
261 *      FDE augmentation section
262 *
263 *
264 *      'L' (and length > 0):
265 *      LSDA            varying      LSDA pointer, encoded in the
266 *      format specified by the
267 *      corresponding operand in the CIE's
268 *      augmentation body.
269 *
270 *      6. Optional Call      varying
271 *      Frame Instructions
272 *
273 * The size of the optional call frame instruction area must be computed
274 * based on the overall size and the offset reached while scanning the
275 * preceding fields of the FDE.
276 *
277 * The overall size of a .eh_frame section is given in the ELF section
278 * header. The only way to determine the number of entries is to scan
279 * the section till the end and count.
280 *
281 */

286 static uint_t
287 extract_uint(const uchar_t *data, uint64_t *ndx, int do_swap)
288 {
289     uint_t r;
290     uchar_t *p = (uchar_t *)&r;

292     data += *ndx;
293     if (do_swap)
294         UL_ASSIGN_BSWAP_WORD(p, data);
295     else
296         UL_ASSIGN_WORD(p, data);

298     (*ndx) += 4;
299     return (r);
300 }

302 /*
303 * Create an unwind header (.eh_frame_hdr) output section.
304 * The section is created and space reserved, but the data
305 * is not copied into place. That is done by a later call
306 * to ld_unwind_populate(), after active relocations have been
307 * processed.
308 *
309 * When GNU linkonce processing is in effect, we can end up in a situation
310 * where the FDEs related to discarded sections remain in the eh_frame
311 * section. Ideally, we would remove these dead entries from eh_frame.
312 * However, that optimization has not yet been implemented. In the current
313 * implementation, the number of dead FDEs cannot be determined until
314 * active relocations are processed, and that processing follows the
315 * call to this function. This means that we are unable to detect dead FDEs
316 * here, and the section created by this routine is sized for maximum case
317 * where all FDEs are valid.
318 */
319 uintptr_t
320 ld_unwind_make_hdr(Of1_desc *of1)
321 {
322     int      bswap = (of1->of1_flags1 & FLG_OF1_ENCDIFF) != 0;
323     Shdr     *shdr;
324     Elf_Data *elfdata;
325     Is_desc  *isp;

```

```

326     size_t   size;
327     Xword    fde_cnt;
328     Aliste   idx1;
329     Os_desc  *osp;

331     /*
332     * we only build a unwind header if we have
333     * some unwind information in the file.
334     */
335     if (of1->of1_unwind == NULL)
336         return (1);

338     /*
339     * Allocate and initialize the Elf_Data structure.
340     */
341     if ((elfdata = libld_calloc(sizeof (Elf_Data), 1)) == NULL)
342         return (S_ERROR);
343     elfdata->d_type = ELF_T_BYTE;
344     elfdata->d_align = ld_targ.t_m.m_word_align;
345     elfdata->d_version = of1->of1_dehdr->e_version;

347     /*
348     * Allocate and initialize the Shdr structure.
349     */
350     if ((shdr = libld_calloc(sizeof (Shdr), 1)) == NULL)
351         return (S_ERROR);
352     shdr->sh_type = ld_targ.t_m.m_sht_unwind;
353     shdr->sh_flags = SHF_ALLOC;
354     shdr->sh_addralign = ld_targ.t_m.m_word_align;
355     shdr->sh_entsize = 0;

357     /*
358     * Allocate and initialize the Is_desc structure.
359     */
360     if ((isp = libld_calloc(1, sizeof (Is_desc))) == NULL)
361         return (S_ERROR);
362     isp->is_name = MSG_ORIG(MSG_SCN_UNWINDHDR);
363     isp->is_shdr = shdr;
364     isp->is_indata = elfdata;

366     if ((of1->of1_unwindhdr = ld_place_section(of1, isp, NULL,
367         ld_targ.t_id.id_unwindhdr, NULL)) == (Os_desc *)S_ERROR)
368         return (S_ERROR);

370     /*
371     * Scan through all of the input Frame information, counting each FDE
372     * that requires an index. Each fde_entry gets a corresponding entry
373     * in the binary search table.
374     */
375     fde_cnt = 0;
376     for (APLIST_TRAVERSE(of1->of1_unwind, idx1, osp)) {
377         Aliste idx2;
378         int    os_isdescs_idx;

380         OS_ISDESCS_TRAVERSE(os_isdescs_idx, osp, idx2, isp) {
381             uchar_t *data;
382             uint64_t off = 0;

384             data = isp->is_indata->d_buf;
385             size = isp->is_indata->d_size;

387             while (off < size) {
388                 uint_t length, id;
389                 uint64_t ndx = 0;

391                 /*

```

```

392         * Extract length in lsb format.  A zero length
393         * indicates that this CIE is a terminator and
394         * that processing for unwind information is
395         * complete.
396         */
397         length = extract_uint(data + off, &ndx, bswap);
398         if (length == 0)
399             break;

401     /*
402     * Extract CIE id in lsb format.
403     */
404     id = extract_uint(data + off, &ndx, bswap);

406     /*
407     * A CIE record has a id of '0', otherwise
408     * this is a FDE entry and the 'id' is the
409     * CIE pointer.
410     */
411     if (id == 0) {
412         uint_t cieversion;
413         /*
414          * The only CIE version supported
415          * is '1' - quick sanity check
416          * here.
417          */
418         cieversion = data[off + ndx];
419         ndx += 1;
420         /* BEGIN CSTYLED */
421         if (cieversion != 1 && cieversion != 3)
422             if (cieversion != 1) {
423                 ld_printf(ofl, ERR_FATAL,
424                     MSG_INTL(MSG_UNW_BADCIEVERS),
425                     isp->is_file->ifl_name,
426                     isp->is_name, off);
427                 return (S_ERROR);
428             }
429         /* END CSTYLED */
430     } else {
431         fde_cnt++;
432     }
433     off += length + 4;
434 }

435 /*
436 * section size:
437 *   byte    version          +1
438 *   byte    eh_frame_ptr_enc +1
439 *   byte    fde_count_enc   +1
440 *   byte    table_enc       +1
441 *   4 bytes eh_frame_ptr    +4
442 *   4 bytes fde_count       +4
443 *   [4 bytes] [4bytes] * fde_count ...
444 */
445 size = 12 + (8 * fde_cnt);
446 if ((elfdata->d_buf = libld_calloc(size, 1)) == NULL)
447     return (S_ERROR);
448 elfdata->d_size = size;
449 shdr->sh_size = (Xword)size;
450 return (1);
451 }

```

unchanged portion omitted

```

482 uintptr_t
483 ld_unwind_populate_hdr(Of1_desc *of1)
484 {
485     uchar_t      *hdrdata;
486     uint_t        binarytable;
487     uint_t        hdroff;
488     Aliste        idx;
489     Addr          hdraddr;
490     Os_desc       *hdrosp;
491     Os_desc       *osp;
492     Os_desc       *first_unwind;
493     uint_t        fde_count;
494     uint_t        *uint_ptr;
495     int           bswap = (of1->ofl_flags1 & FLG_OF1_ENCDIFF) != 0;

497     /*
498     * Are we building the unwind hdr?
499     */
500     if ((hdrosp = of1->ofl_unwindhdr) == 0)
501         return (1);

503     hdrdata = hdrosp->os_outdata->d_buf;
504     hdraddr = hdrosp->os_shdr->sh_addr;
505     hdroff = 0;

507     /*
508     * version == 1
509     */
510     hdrdata[hdroff++] = 1;
511     /*
512     * The encodings are:
513     *
514     * eh_frame_ptr_enc   sdata4 | pcrel
515     * fde_count_enc      udata4
516     * table_enc          sdata4 | datarel
517     */
518     hdrdata[hdroff++] = DW_EH_PE_sdata4 | DW_EH_PE_pcrel;
519     hdrdata[hdroff++] = DW_EH_PE_udata4;
520     hdrdata[hdroff++] = DW_EH_PE_sdata4 | DW_EH_PE_datarel;

522     /*
523     * Header Offsets
524     * -----
525     * byte    version          +1
526     * byte    eh_frame_ptr_enc +1
527     * byte    fde_count_enc   +1
528     * byte    table_enc       +1
529     * 4 bytes eh_frame_ptr    +4
530     * 4 bytes fde_count       +4
531     */
532     /* LINTED */
533     binarytable = (uint_t *) (hdrdata + 12);
534     first_unwind = 0;
535     fde_count = 0;

537     for (APLIST_TRAVERSE(of1->ofl_unwind, idx, osp)) {
538         uchar_t *data;
539         size_t size;
540         uint64_t off = 0;
541         uint_t cieRflag = 0, ciePflag = 0;
542         Shdr *shdr;

544         /*
545         * remember first UNWIND section to
546         * point to in the frame_ptr entry.

```

```

547     */
548     if (first_unwind == 0)
549         first_unwind = osp;

551     data = osp->os_outdata->d_buf;
552     shdr = osp->os_shdr;
553     size = shdr->sh_size;

555     while (off < size) {
556         uint_t    length, id;
557         uint64_t  ndx = 0;

559         /*
560          * Extract length in lsb format. A zero length
561          * indicates that this CIE is a terminator and that
562          * processing of unwind information is complete.
563          */
564         length = extract_uint(data + off, &ndx, bswap);
565         if (length == 0)
566             goto done;

568         /*
569          * Extract CIE id in lsb format.
570          */
571         id = extract_uint(data + off, &ndx, bswap);

573         /*
574          * A CIE record has a id of '0'; otherwise
575          * this is a FDE entry and the 'id' is the
576          * CIE pointer.
577          */
578         if (id == 0) {
579             char    *cieaugstr;
580             uint_t  cieaugndx;
581             uint_t  cieversion;

583             ciePflag = 0;
584             cieRflag = 0;
585             /*
586              * We need to drill through the CIE
587              * to find the Rflag. It's the Rflag
588              * which describes how the FDE code-pointers
589              * are encoded.
590              */

592             cieversion = data[off + ndx];
593             ndx += 1;
594             /*
595              * burn through version
596              */
597             ndx++;

599             /*
600              * augstr
601              */
602             cieaugstr = (char *)&data[off + ndx];
603             ndx += strlen(cieaugstr) + 1;

605             /*
606              * calign & dalign
607              */
608             (void) uleb_extract(&data[off], &ndx);
609             (void) sleb_extract(&data[off], &ndx);

611             /*
612              * retreg

```

```

609     */
610     if (cieversion == 1)
611         ndx++;
612     else
613         (void) uleb_extract(&data[off], &ndx);

614     /*
615      * we walk through the augmentation
616      * section now looking for the Rflag
617      */
618     for (cieaugndx = 0; cieaugstr[cieaugndx];
619          cieaugndx++) {
620         /* BEGIN CSTYLED */
621         switch (cieaugstr[cieaugndx]) {
622         case 'z':
623             /* size */
624             (void) uleb_extract(&data[off],
625                                &ndx);
626             break;
627         case 'P':
628             /* personality */
629             ciePflag = data[off + ndx];
630             ndx++;
631             /*
632              * Just need to extract the
633              * value to move on to the next
634              * field.
635              */
636             (void) dwarf_ehe_extract(
637                 &data[off + ndx],
638                 &ndx, ciePflag,
639                 ofl->ofl_dehdr->e_ident, B_FALSE,
640                 shdr->sh_addr, off + ndx, 0);
641             break;
642         case 'R':
643             /* code encoding */
644             cieRflag = data[off + ndx];
645             ndx++;
646             break;
647         case 'L':
648             /* lsd encoding */
649             ndx++;
650             break;
651         }
652         /* END CSTYLED */
653     }
654 } else {
655     uint_t    bintabndx;
656     uint64_t  initloc;
657     uint64_t  fdeaddr;
658     uint64_t  gotaddr = 0;

660     if (ofl->ofl_osgot != NULL)
661         gotaddr =
662             ofl->ofl_osgot->os_shdr->sh_addr;

664     initloc = dwarf_ehe_extract(&data[off],
665                                &ndx, cieRflag, ofl->ofl_dehdr->e_ident,
666                                B_FALSE,
667                                shdr->sh_addr, off + ndx,
668                                gotaddr);

670     /*
671      * Ignore FDEs with initloc set to 0.
672      * initloc will not be 0 unless this FDE was
673      * abandoned due to GNU linkonce processing.

```

```

674         * The 0 value occurs because we don't resolve
675         * sloppy relocations for unwind header target
676         * sections.
677         */
678         if (initloc != 0) {
679             bintabndx = fde_count * 2;
680             fde_count++;
681
682             /*
683              * FDEaddr is adjusted
684              * to account for the length & id which
685              * have already been consumed.
686              */
687             fdeaddr = shdr->sh_addr + off;
688
689             binarytable[bintabndx] =
690                 (uint_t)(initloc - hdraddr);
691             binarytable[bintabndx + 1] =
692                 (uint_t)(fdeaddr - hdraddr);
693         }
694     }
695
696     /*
697      * the length does not include the length
698      * itself - so account for that too.
699      */
700     off += length + 4;
701 }
702
703 done:
704 /*
705  * Do a quicksort on the binary table. If this is a cross
706  * link from a system with the opposite byte order, xlate
707  * the resulting values into LSB order.
708  */
709     framehdr_addr = hdraddr;
710     qsort((void *)binarytable, (size_t)fde_count,
711         (size_t)(sizeof (uint_t) * 2), bintabcompare);
712     if (bswap) {
713         uint_t *btable = binarytable;
714         uint_t cnt;
715
716         for (cnt = fde_count * 2; cnt-- > 0; btable++)
717             *btable = ld_bswap_Word(*btable);
718     }
719
720     /*
721      * Fill in:
722      *   first_frame_ptr
723      *   fde_count
724      */
725     hdroff = 4;
726     /* LINTED */
727     uint_ptr = (uint_t *)&hdrdata[hdroff];
728     *uint_ptr = first_unwind->os_shdr->sh_addr -
729         (hdrosp->os_shdr->sh_addr + hdroff);
730     if (bswap)
731         *uint_ptr = ld_bswap_Word(*uint_ptr);
732
733     hdroff += 4;
734     /* LINTED */
735     uint_ptr = (uint_t *)&hdrdata[hdroff];
736     *uint_ptr = fde_count;
737     if (bswap)
738         *uint_ptr = ld_bswap_Word(*uint_ptr);
739

```

```

741     /*
742      * If relaxed relocations are active, then there is a chance
743      * that we didn't use all the space reserved for this section.
744      * For details, see the note at head of ld_unwind_make_hdr() above.
745      *
746      * Find the PT_SUNW_UNWIND program header, and change the size values
747      * to the size of the subset of the section that was actually used.
748      */
749     if (ofl->ofl_flags1 & FLG_OF1_RLXREL) {
750         Word phnum = ofl->ofl_nehdr->e_phnum;
751         Phdr *phdr = ofl->ofl_phdr;
752
753         for (; phnum-- > 0; phdr++) {
754             if (phdr->p_type == PT_SUNW_UNWIND) {
755                 phdr->p_memsz = 12 + (8 * fde_count);
756                 phdr->p_filesz = phdr->p_memsz;
757                 break;
758             }
759         }
760     }
761
762     return (1);
763 }

```

unchanged_portion_omitted