## Identification

BCPL-MULTICS comiler - code generation
R. H. Canaday

## Purpose

This document discusses the code produced by the BCPL-MULTICS compiler. Most of this document consists of descriptions of various detai"ed aspects of the code. The reader is presumed to be familiar with Ref(1) and with MSPM sections BZ.6.00 through 3Z.6.03.

This document is not intended for the casual reader. It will be helpful to refer frequently to an EPLBSA listing produced by the BCPL compiler (such as is provided in the appendix) while reading the following discussions.

Summary
This document is subdivided as follows:

## 1) Initialization

Unlike EPL, BCPL programs cannot always be initialized on first reference. For example, a BCPL procedure cannot communicate through the global vector with procedures in other segments unless those segments have been initialized.

## 2) Stack management

Since BCPL does not use a standard MULTICS call-savereturn for intercommunication, a separate BCPL stack is set up and maintained. The programmer has a great deal of (optional) control over the stack and global vector.

## 3) Call-Save-Return

Intercommunication between BCPL procedures is described.
4) Communication with EPL Procedures

Run-time routines have been provided which make possible recursive calls from EPL to BCPL procedures and vice-versa.
5) How to read EPLBSA produced by the BCPL compiler.

The listing from a compilation is not easy to correlate with the soulce code which produced it. However, the details given in this section should make the task somewhat easier.

Initial zation
A BCPL procedure segment name is the same as the name of the CTSS or MULTICS source file which produced it. Unlike EPL , the segment name has nothing to do with the name of any procedure within the segment.

A BCPL wrocedure segment should be initialized before it is executed. The nitialization of a segment named 'XXX' can be done either by calling (from EPL) the initial zation en\%ry woint ${ }^{\prime} X X X \$ X X X$ - [e.g., 'call XXX'] or by calling (from EPL) a procedure within the segment [e.g., "call XXX\$"R02 (arg1, arg2,...)].

Example:
A program consists of three BCPL procedure segments named 'ASEG', 'BSEG', and 'CSEG'. The entry point is the procedure 'PRO3 ${ }^{\circ}$ " $n$ "BSEGं There are no arguments. The EPL driver would be as Follows:

| START: | PROC ( ); |
| :--- | :--- |
|  | CALL ASEG; |
|  | CALL CSEG; |
|  | CALL BSEG\$PR03; |
|  | END; |

In BCPL there are two forms of static storage which must be initialized: "global" and 'local". Local static storage is kept in the linkage segment and is truly static - like EPL static it is nitialized only once in a given process. Repeated calls to initialize it will have no effect. A switch word, "INITSW", in each ilinkage segment is set to zero to indicate that \%nitialization has been done. Global static, on the other hand, is kept in the 'global vector" (see the next section, Global and Stack Management, for details on the allocation of the global vector). Global static can be reinitiolized repeatedly. Calling a BCPL procedure ('call ESEG PRO3") will initialize global static only if it has never been initialized for the segment BSEG (i.e., if IN\%TSW $\neq 0$ ), but calling the initialization entry point ('cal" BSEǴ) will always reinitialize global static for BSEG. This is used to prevent interference between logically separate BCPL programs running in the same process.

Segment 'BCPLGL' contains the routines which initialize stëtic storage and which accept EPL calls to BCPL procedures and do argument conversion. The entry points are:
'BCPLGL\$GINIT' called automatically by BCPL procedures to initialize global static and, if necessary, local static.
'BCPLGL\$GSETU' called automatically by BCPL procedures to do stack setup and argument conversion.

There are also other entry points which are described in later sections of this document.

## Glopal and Stack Management

Since $B C P L$ does not conform to MULTICS standards in its stack menagement, it cannot use "SP' to point to its current stack frame. 'SF' always points to a valid MULTICS stack frame. While BCPL procedures are in execution, 'SP' points to a stack frame immediately inferior to (called by) that of the caller of BCPL, which contains only the header information and two words of diagnostic information: SP|32 contains an ITS pair which identifies the BCPL procedure segment which was called from EPL, and SP|34 contains an ITS pair whict points to the global vector and the stack frame of the first BCPL procedure that was called.

The BCPL global vector and stack are contained in a segment which may be supplied by the caller or routines superior to it, or which will be created during BCPL initialization if not supplied. BCPL uses "unique_chars" and "smm\$set_name_ status" to create a segment of 250K words. The global vector, whict has a standard length of 2048 words, always starts at word 0 of the segment. The stack area begins immediately after the last word of the global vector.

During execution of a BCPL procedure the address register pairs are used as follows:

```
AP - points to current BCPL stack frame
                                    (AB points to global vector)
BP - general usage
SP - MULTICS stack frame
LP - lirkage segment
```

The registers are always paired. $A B$ can be used for the global vector because the global vector always begins at word 0 of the segment containing the BCPL stack frames.

MULTICS SYSTEM-PROGRAMMERS* MANUAL
SECTION BZ.6.04 PAGE 4
BCPL stack segment management centers around a pointer called the "Current Start of Stack" pointer, abbreviated "CSS". This pointrer is kept in static storage and is accessed and changed through four routines. Routine BCPLGL\$SETAP can be called from EPL to set CSS to any address, including null. Routine BCPLGL LOOKAP returns in its argument the current value of CSS. Both entries have one argument, of type "PTR".

The third routine, BCPLGL\$GSETU, is not callable from EPL. It is invoked atitomatically whenever an EPL procedure calls a BCPL procedure. Its function is to do EPL-BCPL argument conversion and to load the address pair AB-AP with the value of CSS.

The fourth routine is the BCPL-callable procedure 'Call', accessed through global vector location 14. Its effect on CSS is to save the old value of CSS and reload CSS with a pointer to the first free word after the current BCPL stack frame. 'Call' then calls the desired EPL procedure. On return the current value of CSS is discarded and the saved value is restored.

The commonest use of these routines is to make sure that two independent BCPL programs (sets of procedures) running in the same process do not conflict in their use of the global vector. The simplest way to guarantee that a BCPL program will be executed with a clean global-and-stack segment is to issue the call

## CALL BCPLGL\$SETAF (NULL)

before initializirg any BCPL procedure segment in the program. Note that any 'CALL BCPLGL\$SETAP' issued after the initialization of a BCPL procedure segment will invalidate that initialization. The only exception to this is that if a BCPL procedure calls an EPL procedure, nothing done by that procedure or any inferior procedure can affect that value of CSS which will be restored when the EPL procedure returns to the BCPL procedure. In fact, a good way of insuring that the global vector will be unchanged by (and invisible to) inferior procedures is to issue the call

> CALL BCPLGL\$SETAP (NULL);
in any EPL procedure called from a BCPL procedure.
In the case of a BCPL-to-EPL-to-BCPL calling chain in which the inferior BCPL procedure should reference the same global vector as the superior one, the superior routine can protect itself from changes made by the inferior routine by issuing the $B C P$ calls:
before calling the EPL routine, and
Restore ( )
on return. The effect of 'Save ( )' and 'Restore ( )' is to silive and restore the global vector in a pushdown stack (which is kept in the MULTICS stack). The call 'Save ( )' does not change the contents of the global vector. 'Save ( )' and 'Restore ( )' can be used not only to bracket a call to an EPL procedure, but in fact at any point in at BCPL procedure.

The routines 'BCF'LGL\$SETSP(PTR)' 'BCPLGL\$LOOKAP(PTR)', 'Save ( )', and 'Restore ( )' give the programmer complete control over global static storage. However, he has very little control over local static storage. Thus in the MULTICS environment it may be useful to use portions of the global vector for static storage. However, be warned that these routines for controlling global static may be dependent on the MULTICS environment. Thus coding dependent on them may not be machine independent. *

## Call-Save-Return

The BCPL stack frame header consists of 8 words, containing the following information:

| $\frac{}{}$ WORD | CONTENTS | STORED BY |
| :--- | :--- | :--- |
| $0-1$ | RETURN ITS SAVE |  |
| $2-3$ | LF FOR CURRENT PROCEDURE | SAVE |
| 4 | UNUSED |  |
| 5 UPPER | LEFT-HAND-SIDE FLAG |  |
| 5 LOWER | ARGUMENT COUNT | CALL |
| $6-7$ | AF FOR SUPERIOR PROCEDURE | CALL |

The code procluced by the compiler for call-save-return in procedure "PROCED" is as follows. In this case the new stack frame will start at the nth word of the previous frame.

CALL

LDA
STA
EAX2
LDA
EABBB
TSBBP
$m, D L$
$A F^{\prime} \mid n+5$
n
CALLADDRESS
O,AU
BÉlO,AP
\# arguments

A BCPL ADDRESS

[^0]MULTICS SYSTEM-PEOGRMMERS' MANUAL


SECTION BZ.6.04 PAGE 6
in the linkage segment

Set LP
The procedure
Save Sequence


RETURN

| EAPAP | AP\| $6, *$ | old AP |
| :--- | :--- | :--- |
| EAPLP | AP $2, *$ | old LP |
| TRA | AP\|O,* | return |

## Communication with EPL Procedures

As has previously been stated, communication between BCPL procedures is by means of a special calling sequence rather than the standard MULTICS call. Some of the significant differences between a BCPL call and a standard MULTICS call are:

1) BCPL arguments do not have dope
2) BCPL arguments are by value.
3) The BCPL stack frame header is 8 words long rather than 32 .
4) Registers are not saved and restored across a call. except for $A B-A P$, $L B-L P$, and $S B-S P$.
5) $\quad S E=-S P$ is not used. The stack pointer is $A B-A P$.
6) The address of the callee is stored in a variable.

The link-by-name features of MULTICS are not used. Obviously, despite these differences, a mechanism must exist to permit calls from EPL procedures to BCPL procedures and vice-versa.

In order to enable EPL procedures to call BCPL procedures the BCPL compiler generates two entry points for each $B C P L$ procedure. One of these is used in BCPL-to-BCPL calls
to the procedure. The other entry point, labeled with the name of the procedure, is callable from EPL. This entry point invokes the routine "BCPLGL\$GSETU' which transforms the cal" into BCPL format and generates a BCPL argument list from the EPL argument list. Dope and specifiers are not interpreted. The addresses in the EPL arglist are translated into BCPL addresses and placed in the BCPL arglist," Dope and specifiers can be interpreted by the callee as desired. Two functions exist to aid in interpreting arguments. These are 'MtoBaddr $(X)$ ' which accepts the (BCPL) address of an ITS pair and returns the value of the ITS pair as a BCPL address, and "MtoBstring (X,V)" which accepts the ( $B C P L$ ) address ' $X$ ' of a MULTICS string specifier and places a corresponding BCPL string in vector ' $V$ '.

Calling from a BCPL procedure to an EPL procedure is done using the two library routines 'Getadr' and 'Call' (cf MSFM $B Z_{" \prime} 6.02$ ). All of the arguments of 'Call' are addresses. The first argument is the address to be called. The remaining adciresses will be converted to ITS pairs and stored in an EPL-format argument list. If any arguments require dope and specifiers, they must be generated by the caller before calling 'Cali". Two library functions, "ITS' and 'BtooMstring", are provided for this purpose. Function 'ITS' creates an ITS pair from a BCPL address. Function 'Bt:oMstring- creates a fixed length MULTICS character string from a BCPL string. MSPM BZ.6.02 contains more detail on 'ITS' and 'BtoMstring'.

The only complicated part of the method for calling between EPL and BCPL procedures is the stack management, which was described in detail above. In general it is possible to call freely between EPL procedures and BCPL procedures if one remembers that in the absence of calls to 'BCPLGL\$SETAP', there will not be any conflicts in stack usage. It is also worth noting that the EPL call

## CALL BCPLGL\$SETAP (NULL);

is always a safe way of guaranteeing that BCPL programs which precede and follow the call will be independent.

## How to read EPLBSA produced by the BCPL compiler

The BCPL compiler produces some comments to help correlate the EPLBSA code with the source program: the source-program name of each variable is given at the point it is declared; the source name of each function and variable labels the corresponding entry point; and each call to a function or routine is labelled, at its occurrence in the code, with the source name of the callee.

Register usage by the code generators is as follows:

| $A$ | $u s e d$ |
| :--- | :--- |
| 0 |  |

AQ used for floating point, multiply, and divide
$\times 3-\times 4$ always used as a pair $-\times 4=$ uppper half word, $\times 3=1$ wer
$x 5-\times 6$ a pair like $\times 3-\times 4$
$\times 2$ contains the stack increment (frame size) during a call
$x 0$ used with $B B$ for indirection
$x 1, x 7$ unused
Multiple location counters are used, as follows:
MAINC procedare segment main counter
LASTC procedure segment builtin subroutines
STATC linkage segment static storage to be initialized

LSTATC linkage segment static storage not needing initialization

IGLøBC procedure segment data to be put in the global vector

The next part of this discussion consists of examples of BCPL statements and the code generated by them, together with a brief description where necessary.

Example A declarations
let $a, b, c=0,0,5$
let $v=\operatorname{vec} 100$

| STZ AP\| $6+10$ | "Declare a |
| :--- | :--- |
| STZ APP $6+11$ | "Declare b |
| LDA 5 DL |  |
| STA AP\| $6+12$ | "Declare c |
| EAPBP AP\| $6+14$ |  |
| EAX3 AP $6+14$ |  |
| SBRBB AP\| $6+13$ |  |
| SXL3 AP\| $6+13$ | "Declare v |

This is an example of a very common operation, namely add'ress generation. The variable ' $v$ " is at location 19 of the stack frame. The vector begins at location 20. '19' is written ' $6+13^{\prime}$ because of a historical carry-over from GMiP.

Example B
$a:=v[20]$

## assignments

| LDA | AP $\mid 6+13$ | $v$ |
| :--- | :--- | :--- |
| EABBB $O, A U$ |  |  |
| LDA | BB $\mid 20, A L$ |  |
| STA | AP\| $6+10$ | $a$ |

This is an example of using an address. Addresses may be used from registers $A$ or $Q$, or from register pairs $\times 3-\times 4$ or $\times 5-\times 6$.

$$
v[20]:=b[a]
$$

## Examplec

$v[0]:=v[20] * a[b+c]$

| LDA | AP\|6+10 | a |
| :---: | :---: | :---: |
| ADA | AP $\mid 6+11$ | b |
| ADA | AP\| $6+12$ | c |
| EAX4 | O,AU |  |
| EAX3 | O,AL |  |
| LDQ | $A P \mid 6+13$ | v |
| EABBB | O,QU |  |
| LDQ | BB\|20,QL | v[20] |
| EABBB | 0,4 |  |
| MPY | BB\|0,3 | $a[b+c]$ |
| EABBB | AP $\mid 6+13$ | $v$ |
| STA | B $\mid 0,8$ |  |

## Example D

astring (does not need initialization)

L5: NULL
DCT
ФCT

LSTATC
005141142143
144145000000
an address (this will be transformed into a BCPL address which will be stored in place of the EAPBP)

US STATC
L6: $N U_{i}^{\prime} L$
EAPBP LP|L8 label in linkage seg.
L7: NU L
EAPBP
AP|L9 label in procedure seg.
an address to go in global vector location 20
$\begin{array}{lcl}\text { USE } & \text { IGLØBC } \\ \text { EAPBP } & A B \mid L 10 & \text { label in procedure seg. } \\ \text { ZEFO } & 0,20 & \end{array}$

```
3CPL invoked through Mrgede
Oontrol Cafd = old bepl testg9
// This is a demonstration of BCPL code generation,
// The prögram doe: not do anything and cannot be executed.
global $l Demonstration;100; Functions101 $)
let Demonstration (A1,A2) * Yalos $!
    let a,b;c=0
    let v = vec 100
    a := v[20]
    v[20] i= i[a]
    v[0] := v[20] a[b+c]
    c i= "This is a string"
    a := Function(a, A1)
    Label:
        Label := Label + A2
        resultls A" / label $)
```

APPENDIX





## OUTPUT PROM EPLBSA ASSEMBLy

| AA | 000101 | 0 | 00010 | 236100 | 91 |  | IDQ | $A P * 6 * 2$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Q ${ }^{\text {a }}$ | 000102 | 4 | 00014 | 506900 | 92 |  | DIV | 1P* 15 |
| AA | 000103 |  | 00000 | 756900 | 93 |  | STQ | $\mathrm{AB}^{+} \mathrm{O}$ |
| OA | 000104 |  | 000105 | 790000 | 94 |  | tra | 28 |
|  |  |  |  |  | 95 |  | USE | ISTATC |
|  | 000022 |  |  |  | 96 | L6: | NULL |  |
| Aa | 000022 |  | 020124 | 150951 | 97 |  | OCT | 020124950151 |
| AA | 000023 |  | 163040 | 151963 | 98 |  | оет | 163040951963 |
| \% | 000024 |  | 040141 | 040163 | 99 |  | OCT | OH0445040156 |
| AA | 000025 |  | 164162 | 151956 | 100 |  | OCT | 164162951156 |
| an | 000026 |  | 147000 | 000000 | $\begin{aligned} & 101 \\ & 102 \end{aligned}$ |  | OCT USE | 147000000000 STATC |
|  | 000014 |  |  |  | 103 | 25\% | N0LL |  |
| AA | 000014 | 1 | 00076 | 352100 | 104 |  | eapbi | $A^{\text {B }} \mathrm{E}$ [ 7 |
|  |  |  |  |  | 105 |  | USE | MAINC |
|  | 000105 |  |  |  | 106 | 18: | NULI |  |
| Aa | 000105 | 0 | 00006 | 350120 | 107 |  | eapap | AP* 6 , ${ }^{\text {a }}$ |
| AA | 000106 |  | 00002 | 370120 | 108 |  | EAPIP | AP*2,* |
| Aa | 000107 | 0 | 00000 | 610100 | 109 |  | RTCD | AP*O |
|  | 000110 |  |  |  | 190 | 13: | NULI, |  |
|  | 000910 |  |  |  | 199 |  |  |  |
|  | 000190 |  |  |  | 112 | " |  |  |
|  | 000110 |  |  |  | 193 | " Max | STACK USED | 128 |
|  |  |  |  |  | 994 |  | USE | IGIOBC |
|  | 000110 |  |  |  | 915 | 22: | NULI |  |
| 4 A | 000110 | 4 | 00097 | 352100 | 996 |  | EAPbp | 12* 14 |
| AA | 000119 |  | 000000 | 000144 | 117 |  | 2ERO | 0.100 |
| AA | 000112 |  | 000000 | 000000 | 198 |  | ZERO |  |
| Aa | 000113 |  | 000000 | 000000 | 199 |  | 2ERO |  |
|  |  |  |  |  | 120 |  | USE | Statc |
| A 1 | 000015 |  | 000000 | 352000 | 121 |  | EAPB? |  |
|  |  |  |  |  | 122 |  | USE | LASTC |
|  | 000914 |  |  |  | 123 | Sv: | NULE |  |
| AA | 000194 | 0 | 00006 | 250112 | 124 |  | Stpap | $A P^{*} 6,2$ |
| Aa | 000115 | 0 | 00000 | 350912 | 125 |  | EAPAP | $A P^{*} 0.2$ |
| Aa | 000116 | 0 | 00002 | 6509 00 | 126 |  | STPLP | $A P^{*} 2$ |
| 4a | 000119 | 4 | 00016 | 234900 | 127 |  | SEN | LP*EIESW |
| AA | 000120 |  | 000000 | 600017 | 128 |  | T2E | 0.7 |
|  | 000121 |  |  |  | 129 | IMIT8 | NuLE |  |


| 42000129 | 4000162 | 235900 | 130 | LDA | LP＊INITSM |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 41000122 | 400016 | 450900 | 131 | SI | LP＊INITSW |  |  |  |
| AA 000123 | 600000 | 650900 | 132 | STPLP | SP＊ |  |  |  |
| OA 000124 | 000910 | 626000 | 133 | EAX6 | IGIOB |  |  |  |
| 2A 000125 | 000014 | 625000 | 134 | EAX5 | STAT＋4 |  |  |  |
| 41 500126 | 400050 | 272920 | 135 | TSBBP | 〈BCPIGL＞＊${ }^{\text {cINIT }}$ ¢ | RETURA By | TRA | BB＊ 0.9 |
| 000127 |  |  | ${ }^{153}$ | E\％EETV：\％uEz |  |  |  |  |
| Ah 000127 | 600022 | 352920 | 137 | SAVE |  |  |  |  |
| AA 000130 | 200020 | 652100 |  |  |  |  | $\stackrel{3}{\text { ² }}$ |  |
| AA 000131 | 200040 | 3529 6̂ |  |  |  |  |  |  |
| an 000132 | 277762 | 252100 |  |  |  |  |  |  |
| A1 000133 | 277740 | 332400 |  |  |  |  | $\sigma$ |  |
| 5i 5000134 | ¢ 00032 | 250900 |  |  |  |  |  |  |
| OA 000135 | 000194 | 352000 | 138 | EAPBP． | SY |  | $\frac{2}{6}$ |  |
| A1 000136 | 600000 | 650100 | 499 | STPEP | SPO |  | $\square$ |  |
| 4A 000137 | 4.00052 | 710120 | $\begin{aligned} & 140 \\ & 949 \end{aligned}$ | $\begin{aligned} & \operatorname{TRA} \\ & \text { END } \end{aligned}$ | 〈BCPIGL）＂［GSETU］ |  | $\frac{\pi}{3}$ |  |
|  |  |  |  |  |  |  | ㄲ |  |
| NO LItera |  |  |  |  |  |  | ¢ |  |
| entay poi | TS AND SE | EGDEF N |  |  |  |  | $\frac{8}{8}$ |  |
| $\begin{aligned} & 5 A 000140 \\ & 2 A 0009149 \end{aligned}$ | 000006 000044 | $\begin{aligned} & 000000 \\ & 000009 \end{aligned}$ |  |  |  |  | 尓 |  |
| A1 000142 | 095104 | 145155 | 142 | demonstration |  |  | $\sim$ |  |
| A1 000143 | 957 156 | 163164 |  |  |  |  |  |  |
| A1 000144 | 162944 | 164159 |  |  |  |  | $\frac{3}{1}$ |  |
| AA 000145 | 157156 | 000000 |  |  |  |  | z |  |
| 51 000146 | 000012 | 000000 |  |  |  |  | 5 |  |
| 2A 000147 | 000036 | 000001 |  |  |  |  |  |  |
| 14000150 | 006164 | 145163 | 143 | TEST99 |  |  |  |  |
| AA 000159 | 164079 | 071000 |  |  |  |  | m |  |
| 51000952 | 000020 | 000000 |  |  |  |  | $\bigcirc$ |  |
| 6A 000153 | 000000 | 000002 |  |  |  |  | $\bigcirc$ |  |
| AA 000154 | 014163 | 179155 | 144 | SYMBOLTTABLE |  |  | 2 |  |
| AA 000155 | 142157 | 454 137 |  |  |  |  | O |  |
| A1 000156 | 164949 | 142154 |  |  |  |  | N |  |
| AA 000157 | 145000 | 000000 |  |  |  |  | or |  |
| 51 000160 | 000025 | 000000 |  |  |  |  | － |  |
| 61000969 | 000030 | ，000002 |  |  |  |  |  |  |
| AA 000162 | 010162 | 145 154 | 145 | REITEXT |  |  |  |  |
| A1 000163 | 137964 | 145170 |  |  |  |  | 刀 |  |
| 11000164 | 164000 | 000000 |  |  |  |  | $\cdots$ |  |
| 51000165 | 000032 | 000000 |  |  |  |  |  |  |
| ） |  |  |  | ） |  |  | 0 |  |




[^0]:    * The CTSS (and GECOS) environments act as if a 'BCPLGL\$SETAP(NULL)' call were issued at every return to command level ('Jetween activities).

