z/TPF EE V1.1
z/TPFDF V1.1
TPF Toolkit for WebSphere® Studio V3
TPF Operations Server V1.2

IBM

IBM Software Group

TPF Users Group Fall 2005

IBM2047 Code Page

Name : Robert Kopac
Venue : Hot Topics

AIM Enterprise Platform Software

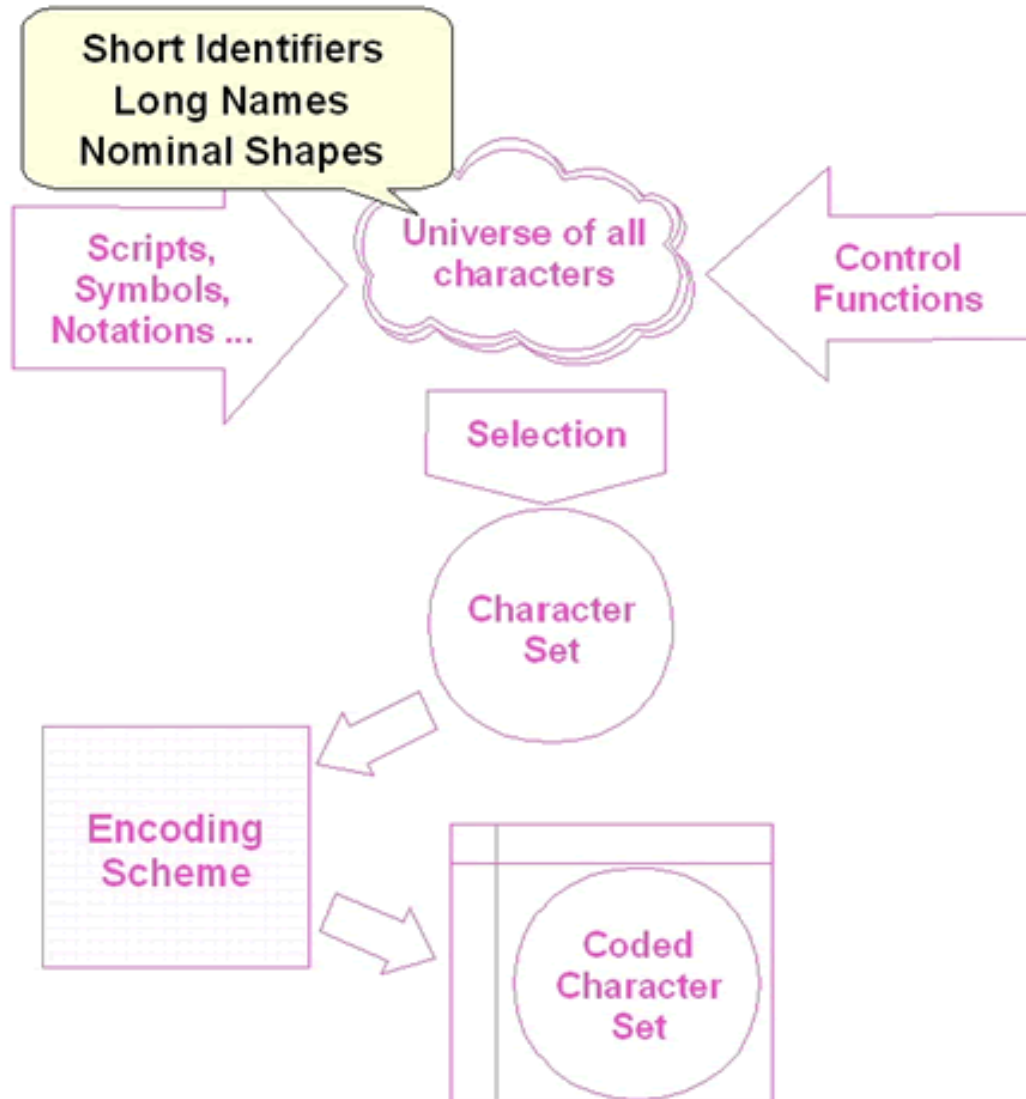IBM z/Transaction Processing Facility Enterprise Edition 1.1.0

© IBM Corporation 2005

z/TPF® Problem to be Solved

- For IBM z/TPF, need the ability to use a non-standard code page when compiling from the command line
  - ► This code page must define linefeed to be x'15'
    - − This is the linefeed value for 3215 terminal devices

- Requires action by customer's Linux® administrator

- Otherwise, compiles from the command line of C or C++ code using linefeed such as '\n' will generate error messages similar to

```
cc1: conversion from cpv to IBM2047 not supported by iconv
```

Background: Encoding schemes and Coded Character Sets

- A Coded Character Set or Code Page is a mapping of a set of numbers (code points) to the corresponding characters assigned to each one.

- an IBM EBCDIC code page can map x'C1' to letter 'A'.

- Coded Character Sets must follow a set of encoding rules, defined by an Encoding Scheme that specifies:
  - ► Form and values for numbers and bit patterns
  - ► Rules for assigning characters to bit patterns

- Control Characters (for example, linefeed) are specified by, and associated with, an Encoding Scheme definition.

Short Identifiers
Long Names
Nominal Shapes

Scripts,
Symbols,
Notations ...

Universe of all
characters

Control
Functions

Selection

Character
Set

Encoding
Scheme

Coded
Character
Set

ISO IBM1047 Code Page and the Line Feed problem

- The International Organization for Standardization (ISO)® created standards for "POSIX® standard" systems
  - ► Including coded character sets
  - ► ISO IBM1047 coded character set most closely matches the characters that z/TPF uses
    - But not exactly!
    - ISO IBM1047 Defines line feed to be x'25'
      - Part of the EBCDIC-2 encoding scheme definition
    - However, z/TPF uses x'15' for linefeed
      - Used with older 3215 terminal devices

Why TPF4.1® has no Code Page Problem

- TPF 4.1 has no code page problem because it uses
  - ► z/OS C/C++ compiler, prelinker, linker
  - ► *Non-standard* z/OS CP1047 code page
    - – with x'15' linefeed

Why z/TPF has a Code Page Problem

- z/TPF has a code page problem because it uses
  - ► GNU compiler collection (GCC)®
    - tpf-gcc is the cross-compiler that allows users to build code on linux systems and load the executable objects on their z/TPF systems.
  - ► GNU C library (glibc)
  - ► GNU standard C++ library (libstdc++)
- Generic character set conversion
  - ► handled by glibc iconv(), also known as "gconv" support
  - ► uses shared loadable modules, including ISO IBM1047
- ISO IBM1047.so is built and shipped as part of glibc
  - ► in /usr/lib or in /usr/lib64

Why z/TPF Cannot Modify IBM1047

- The GCC cannot simply build and ship a new code page that maps linefeed to x'15'
  - ► Code page must be consistent with Encoding Scheme.

- What if z/TPF modifies IBM1047.so to map linefeed to x'15'?
  - ► Problem: If any other Linux application, for example, APACHE HTTPD®, requires the original ISO IBM1047 and handles the linefeed swap with procedural code, a patch to IBM1047.so might break that application.

A tpf-gcc Solution -- IBM2047.so

- Since gconv support is extendable:
  - ► z/TPF created IBM2047.so
    - A patched code page with x'15' linefeed
    - Not an "official" code page name
  - ► z/TPF extended the gconv support
  - ► z/TPF updated the maketpf build tool to recognize it

  - ► z/TPF customers may need to update their Linux system configuration
    - To allow users to run tpf-gcc from command line and use IBM2047.so

How z/TPF Extended gconv

- z/TPF:
  - ► Created a text file called **gconv-modules**
    - – Identifies the new code page translation module IBM2047.so and the context in which it is to operate
  - ► Placed this file in a subdirectory `${TPFROOT}/linux/gconv`
  - ► Copied the new IBM2047.so into the same subdirectory as **gconv_modules**
- The z/TPF installation process uses `${TPFROOT}/linux/gconv`
- z/TPF updated maketpf to use this path

Environment Variable GCONV_PATH

- **maketpf sets an internal variable, not environment variable GCONV_PATH, to point to the z/TPF extension subdirectory**
  - ► However, this does not work if, for example, tpf-gcc is run from the command line
  - ► Without the GCONV_PATH extension and its parts in place, tpf-gcc will fail to run to completion when `-fexec-charset=IBM2047` is passed on the command line
    - – Will get error message similar to

```
 cc1: conversion from cpv to IBM2047 not supported by
iconv
```

- **Solution: set GCONV_PATH to the directory maketpf uses:** `${TPFROOT}/linux/gconv`

Environment Variable GCONV_PATH

- At each login, GCONV_PATH can be set to whatever a user or system administrator would like.

- `/etc/profile` is the file read by all login shells to establish system preferences and rules for shell sessions

- One major role of this script is to possibly set, clear, or extend environment variable settings

- Besides `/etc/profile`, other script files might be read and executed which complete the process of shell initialization

- bash users should look at `${HOME}/.bash_profile`

- For other shells, see your manual pages

- Issue `echo $GCONV_PATH` to view current settings

Customer Action

- For GNU/Linux systems on which tpf-gcc is expected to emit proper code, the GNU/Linux system adminstrator should:
  - ► Where the variable expansion `${TPF_ROOT}` expands to the file system path at which z/TPF source code can be permanently found, add the following lines (first is an optional example) anywhere (toward the end, preferred) of file `/etc/profile`:

    ```
    TPF_ROOT=/my/path/to/the/base/source/installation
    GCONV_PATH=${TPF_ROOT}/linux/gconv:${GCONV_PATH}

    export GCONV_PATH
    ```

- If the system administrator does not perform this modification, individual users should alter their own `$HOME/.bash_profile` files

Conclusion

- After performing these modifications, users when compiling using the command line:
  - ► Will receive correct linefeed characters when specified in C or C++ source code; for example, '\n'
  - ► Will not receive error or warning messages for translation to IBM2047 code page when the `-fexec-charset=IBM2047` option is present on the command line or in compiler default values.

Trademarks

IBM, z/TPF, and TPF4.1 are trademarks of International Business Machines Corporation in the United States, other countries, or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

GNU Compiler Collection, GCC, and GNU are trademarks of Free Software Foundation, Inc. in the United States, other countries, or both.

APACHE HTTPD is a trademark of The Apache Software Foundation in the United States, other countries, or both.

ISO is a trademark of the International Organization for Standardization in the United States, other countries, or both.

POSIX is a trademark of the IEEE Standards Association in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Notes

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.