

OpenSolaris 3分クッキング

☆レシピ 第4巻☆

DTrace ToolKitの紹介 編

Solaris 3分クッキング: DTrace ToolKit の紹介

OpenSolaris DTrace プロジェクトからは、システム管理や DTrace D 言語プログラミングの学習にも役に立つ DTrace ToolKit が提供されています。今夜は、そのツールキットの内容を紹介します。

HotTopics セミナーでは、10 分程度の紹介なので、ほんのさわりだけです。DTrace Toolkit に興味がある方は、こちらのレシピを参考にしてください。

第 1 章: 最新の DTrace Toolkit の入手とインストール

1-1: DTrace Toolkit の動作環境

DTrace Toolkit は、Solaris 10 以降、OpenSolaris (SXDE or SXCE), OpenSolaris 2008.05 上で動作します。プロセッサのアーキテクチャには依存しないので、SPARC と x86 両方で動作します。

1-2: DTrace Toolkit のダウンロードと展開

Toolkit の入手方法とインストール方法を紹介します。DTrace Toolkit は、<http://www.opensolaris.org/os/community/dtrace/dtrac toolkit/> で紹介されています。このページの `DTraceToolkit-0.99.tar.gz` をダウンロード、もしくは以下の方法でダウンロードします。なお、OpenSolaris 2008.05 をご使用の場合は、`ver.0.99` が、`/opt/DTT` の既にインストールされているのでダウンロードとインストールは不要です。

```
hara@opensolaris:~$ wget http://www.brendangregg.com/DTraceToolkit-0.99.tar.gz
--21:19:16-- http://www.brendangregg.com/DTraceToolkit-0.99.tar.gz
=> `DTraceToolkit-0.99.tar.gz'
Resolving www.brendangregg.com... 64.202.163.192
Connecting to www.brendangregg.com|64.202.163.192|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 474,697 (464K) [application/x-tar]

100%[=====>] 474,697      175.97K/s

21:19:19 (175.46 KB/s) - `DTraceToolkit-0.99.tar.gz' saved [474697/474697]

hara@opensolaris:~$
```

`DTraceToolkit-0.99.tar.gz` から `DTraceToolkit-0.99` というディレクトリの下に全てのツールが展開されて置かれます。

```
hara@opensolaris:~$ gunzip DTraceToolkit-0.99.tar.gz
hara@opensolaris:~$ tar -xf DTraceToolkit-0.99.tar
```

さらに、 DTraceToolkit-0.99 の下の install スクリプトを実行することで、 /opt/DTT の下に Toolkit をインストールできます。

```
# ./DTraceToolkit-0.99/install
DTraceToolkit Installation
-----
DTraceToolkit version 0.99, 30-Sep-2007

hit Ctrl-C any time you wish to quit.

Enter target directory for installation [/opt/DTT]:
```

1-3: DTrace Toolkit 利用のための設定

DTrace は、デフォルトは root ユーザの権限でのみ動作します。ユーザモードで実行する場合は、DTrace 実行するための権限をそのユーザに与えます。

```
# usermod -k defaultpriv=basic,dtrace_proc,dtrace_kernel hara <-- 設定するユーザ名
```

コマンド検索パスや、マンパスにも追加しておく便利です。

```
hara@opensolaris:~$ grep PATH ~/.bashrc
export PATH=/usr/gnu/bin:/usr/bin:/usr/X11/bin:/usr/sbin:/sbin:/opt/DTT/Bin
export MANPATH=/usr/gnu/share/man:/usr/share/man:/usr/X11/share/man:/opt/DTT/Man
hara@opensolaris:~$
```

OpenSolaris 2008.05 で使用の場合は、pfexec(1) で、dtrace の使用も可能です。

```
hara@opensolaris:~$ pfexec dtrace -l
```

第 2 章: ツールの概要

展開した DTraceToolkit-0.99 ディレクトリ、あるいは /opt/DTT ディレクトリ以下の各ディレクトリは、以下のような構成になっています。

ディレクトリ名	概要
Bin/	全てのスクリプトへのシンボリックリンク
Apps/	アプリケーション固有のスクリプト・サンプル
Cpu/	CPU 解析用のスクリプト
Code/	練習用サンプルコード
Disk/	ディスク I/O 解析用スクリプト
Docs	このツールキットのドキュメント
Examples/	各スクリプトの使用例集
Include/	D スクリプト用 include ファイル
Java/	Java トレース用スクリプト・サンプル
JavaScript/	JavaScript トレース用スクリプト・サンプル

Kernel/	カーネル解析用スクリプト・サンプル
Locks/	ロック解析用スクリプト・サンプル
Man/	DTrace ツールキットの man ページ
Mem/	メモリ解析用スクリプト・サンプル
Misc/	その他のスクリプト
Net/	ネットワーク解析用スクリプト・サンプル
Notes/	DTrace ツールキットのノート
Perl/	Perl トレース用スクリプト・サンプル
Php/	PHP トレース用スクリプト・サンプル
Proc/	プロセス解析用スクリプト・サンプル
Python/	Python トレース用スクリプト・サンプル
Ruby/	Ruby トレース用スクリプト・サンプル
Shell/	シェル言語トレース用スクリプト・サンプル
Snippets/	D 言語用材料集
System/	システム解析用スクリプト・サンプル
Tcl/	Tcl トレース用スクリプト・サンプル
User/	ユーザ・ベースの活動分析用スクリプト・サンプル
Zones/	Zone 解析用スクリプト・サンプル

特に便利で、よく使われる DTrace ツールは /opt/DTT/ ディレクトリ下に直接置かれています。概要は以下のとおりです。

ツール名	概要
dexplorer	system explorer の DTrace 版、system 情報のコレクションを取って指定したディレクトリに gzip アーカイブファイルとして記録します。アーカイブファイルはシステム診断の情報として利用されます
dtruss	DTrace 版の truss(1)、内容は truss(1) 簡易版ですが、D スクリプトの参考として有効です
dvmstat	DTrace 版の vmstat(1M)、-p PID 指定で PID 毎の VM 情報がとれます
errinfo	system の error message を表示します
execsnoop	このコマンド起動以降にシステム上で起動された実行プログラムを表示します
hotkernel	カーネルレベルの function あるいは module の使用頻度を表示します
hotuser	特定 PID におけるユーザレベルの function または library の使用頻度の表示します
iopattern	Disk の I/O アクセスのパターンを出力します
iosnoop	Disk I/O アクセスの全てのイベントを出力します
iotop	プロセス毎の Disk I/O イベントを表示します
opensnoop	このコマンド起動以降に open されたファイルを全て表示。特定ファイルの監視も可能です
procsystime	あるプロセスの全てのシステムコール時間を表示します
rwsnoop	全ての read/write イベントを表示します
rwtop	プロセス毎の read/write イベントを表示します
statsnoop	file stat の snoop を行います

第3章：ツールの実行例

200 例以上あるツールの中から、いくつかの実行例を紹介します。

ツール名: dexplorer

DTrace スクリプト・コレクションの実行とシステム情報の収集

オプション:

- q 非出力モードの指定
- y 全ての問い合わせに対し "yes" の応答とみなします
- D 出力用のディレクトリを削除しない
- T tar.gz のアーカイブ・ファイルを作成しない
- d <outputdir> 出力先のディレクトリの指定
- i <interval> サンプルの間隔(秒)の指定 (* 指定しない場合: 5 秒間隔)

実行例:

```
$ dexplorer
Output dir will be the current dir (/export/home/hara).
Hit enter for yes, or type path:
Starting dexplorer ver 0.76.
Sample interval is 5 seconds. Total run is > 100 seconds.
 0% Interrupts by CPU...
 5% Interrupt times...
10% Dispatcher queue length by CPU...
[...]
95% Syscall error counts...
100% Done.
File is de_opensolaris_200806241450.tar.gz
$
```

ツール名: dtruss

DTrace 版の truss(1)

```
hara@opensolaris:~$ dtruss -p `pgrep soffice.bin`
SYSCALL(args) = return
ioctl(0x7, 0x4004667F, 0x804746C) = 0 0
pollsys(0x83345F0, 0x7, 0x80474F8) = 0 0
ioctl(0x7, 0x4004667F, 0x804746C) = 0 0
pollsys(0x83345F0, 0x7, 0x80474F8) = 0 0
ioctl(0x7, 0x4004667F, 0x804746C) = 0 0
[...]
```

ツール名: iosnoop

ディスク I/O イベントの詮索/スヌープ

```
hara@opensolaris:~$ iosnoop
UID  PID  D  BLOCK  SIZE  COMM  PATHNAME
101  870  R  3014912 131072 soffice.bin <none>
101  870  R  3015168 131072 soffice.bin <none>
0    0    W  222619136 131072 sched <none>
[...]
```

ツール名: execsnoop

システム上で起動された実行プログラムの表示

```
hara@opensolaris:~$ execsnoop
UID  PID  PPID  ARGS
101  973  971  find . -name opensolaris -print
0    974  971  su
0    978  977  vi /etc/passwd
0    979  977  cat /etc/hosts
101  985  971  more /etc/nsswitch.conf
```

ツール名: hotkernel

カーネル上のファンクションもしくはモジュール(-m) の使用頻度の表示

```
hara@opensolaris:~$ hotkernel -m
Sampling... Hit Ctrl-C to end.
^C
MODULE                                COUNT  PCNT
dtrace                                 1      0.0%
zfs                                     1      0.0%
ehci                                    1      0.0%
hdaudio                                4      0.1%
sockfs                                  21     0.4%
fifofs                                  38     0.6%
osscommon                              408    6.9%
genunix                                690   11.6%
unix                                    4760  80.4%
```

ツール名: opensnoop

open されたファイルの表示

```
hara@opensolaris:~$ opensnoop
UID  PID  COMM      FD  PATH
0    843  nscd      5   /etc/svc/volatile/repository_door
0    843  nscd      5   /etc/svc/volatile/repository_door
0    843  nscd      5   /etc/svc/volatile/repository_door
0    843  nscd      5   /etc/svc/volatile/repository_door
0    843  nscd      5   /etc/inet/ipnodes
0    843  nscd      5   /etc/resolv.conf
0    843  nscd     -1   /etc/hesiod.conf
0    843  nscd      5   /etc/resolv.conf
101  1021  vim       4   /etc/passwd
101  1021  vim       4   /etc/passwd
```

ツール名: rwsnoop

システムコール read/write イベントの詮索/スヌープ

```
hara@opensolaris:~$ rwsnoop -p `pgrep nscd`
UID  PID  CMD      D  BYTES  FILE
0    843  nscd     R  12627  /etc/security/prof_attr
0    843  nscd     R  12627  /etc/security/prof_attr
0    843  nscd     R  12627  /etc/security/prof_attr
0    843  nscd     R   1074  /etc/inet/hosts
0    843  nscd     R     0   /etc/inet/hosts
0    843  nscd     R    20   /etc/resolv.conf
0    843  nscd     R     0   /etc/resolv.conf
0    843  nscd     R   1074  /etc/inet/hosts
```

ツール: swapinfo.d

仮想メモリ(swap)情報の表示

```
hara@opensolaris:~$ swapinfo.d
RAM _____ Total 3062 MB
RAM      Unusable   8 MB
RAM      Kernel    894 MB
RAM      Locked     0 MB
RAM      Used       392 MB
RAM      Free       1766 MB

Disk _____ Total 2047 MB
Disk      Resv      290 MB
Disk      Avail    1756 MB

Swap _____ Total 3825 MB
Swap      Resv      290 MB
Swap      Avail    3534 MB
Swap      (Minfree) 381 MB
```

ツール: tcpstat.d

TCP 統計情報の表示

```
hara@opensolaris:~$ tcpstat.d
TCP_out TCP_outRe TCP_in TCP_inDup TCP_inUn
0 0 0 0 0
133 0 0 0 0
0 0 71719 0 0
0 0 128340 0 0
0 0 111780 0 16560
0 0 35880 0 5520
0 0 35880 0 13800
0 0 46920 0 0
0 0 5849 0 0
0 0 0 0 0
```

ツール名: crash.d

クラッシュダンプのトレース

```
hara@opensolaris:~$ crash.d
Waiting for crashing applications...

-----
CRASH DETECTED at 2008 Jun 30 16:32:53
-----

Type: SIGSEGV
Program: a.out
Args: ./a.out%0
PID: 1309
TID: 1
LWPs: 1
PPID: 968
UID: 101
GID: 10
TaskID: 68
ProjID: 10
PoolID: 0
ZoneID: 0
zone: global
CWD: /export/home/hara
errno: 0
[...]
```

ツール名: j_flow.d

Java メソッドのトレース

注) Java プログラム実行時 `-XX:+ExtendedDTraceProbes` オプションを付ける

```
hara@opensolaris:~$ cat TestGreeting.java
public class TestGreeting {
    public static void main(String[] args) {
        Greeting hello = new Greeting();
        while (true) {
            hello.greet();
            try {
                Thread.currentThread().sleep(1000);
            } catch (InterruptedException e) {
            }
        }
    }
}

hara@opensolaris:~$ javac TestGreeting.java
hara@opensolaris:~$ java -XX:+ExtendedDTraceProbes TestGreeting
Hello DTrace!
Hello DTrace!
Hello DTrace!
Hello DTrace!
[...]
hara@opensolaris:~$
hara@opensolaris:~$ /usr/java/bin/jps -l
1417 sun.tools.jps.Jps
1416 TestGreeting

hara@opensolaris:~$ j_flow.d -p 1416

C   PID TIME(us)      -- CLASS.METHOD
0   1416 19557739254      <- java/lang/Thread.sleep
0   1416 19557739277      -> Greeting.greet
0   1416 19557739287      -> java/io/PrintStream.println
0   1416 19557739298      -> java/io/PrintStream.print
0   1416 19557739305      -> java/io/PrintStream.write
0   1416 19557739312      -> java/io/PrintStream.ensureOpen
0   1416 19557739319      <- java/io/PrintStream.ensureOpen
0   1416 19557739326      -> java/io/Writer.write
0   1416 19557739334      -> java/io/BufferedWriter.write
0   1416 19557739342      -> java/io/BufferedWriter.ensureOpen
0   1416 19557739349      <- java/io/BufferedWriter.ensureOpen
0   1416 19557739357      -> java/io/BufferedWriter.min
0   1416 19557739365      <- java/io/BufferedWriter.min
0   1416 19557739371      -> java/lang/String.getChars
0   1416 19557739380      -> java/lang/System.arraycopy
0   1416 19557739389      <- java/lang/System.arraycopy
0   1416 19557739402      <- java/lang/String.getChars
0   1416 19557739410      <- java/io/BufferedWriter.write
0   1416 19557739416      <- java/io/Writer.write
[...]
```

<作成: サン・マイクロシステムズ 原口>

参考 URL

Brendan Gregg's DTrace Tools

<http://www.brendangregg.com/dtrace.html>

DTrace Topics from SolarisInternals.com

http://www.solarisinternals.com/wiki/index.php/DTrace_Topics