RedHawk6.0 の特徴

RHEL 6.0 をベースに、多くの公式アップデートを、あらかじめ適用しています。
バージョンアップに伴い、更新された主要なパッケージは、以下の通りです。

- GCC 4.4.5 (vs. 4.1.2)
- Glibc2.12 (vs. 2.5)
- Firefox 3.6 (vs. 3.0)
- OpenOffice3.2 (vs. 2.3)
- Gnome 2.28 (vs. 2.16)
- KDE 4.3 (vs. 3.5)
- XorgServer 1.7.7 (vs. 1.1.1)
- Emacs23.1.1 (vs. 21.4.1)
- Vim 7.2.411 (vs. 7.0.109)

RedHawk6.0 カーネルは Kernel.org 2.6.36 をベースに、リアルタイム化したものです。
多くのコミュニティによって拡張された Linux 2.6.36 カーネルには、以下の機能が提供されています

- 新しいパフォーマンスイベント（ブランチ追跡）
- マルチコアに最適化された TCP/IP スタック（スループット向上）
- 多数の仮想化拡張
  - カーネル同一ページのマージング（ksmd）
  - インタラプト・インジェクション（irqfd）
  - 理論速度と同等なネットワーキング（VhostNet）

それに加えて、RedHawk6.0 には、若干のコミュニティリアルタイム OS パッチと、多くのコンカレントリアルタイム OS パッチが含まれ、以下のリアルタイム新機能が提供されています。

- mlockall_pid()
- RAM ディスクシールディング
- NUMA シールディング
- 新しいスケジューリングポリシー
- NUMA CPU トポロジーを見る新しい run コマンドオプション

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

RedHawk6.0 のインストール手順

- RHEL6.0 インストール
  - 新しい単純化された、そしてより速いインストール 。
  - 自動的にディスク（LVM）を分割させることができます。
  - 自動的にスワッピングスペースを設定させることができます。
  - 手作業で「すべてのパッケージを選択する」必要はありません
  - “ソフトウェア開発ワークステーション”を選んでください
  - 手作業でネットワーク をイネーブルにしなくてはなりません
  - SELinux , Firewall をディセーブルにする必要は、ありません。
    - RedHawk インストレーション時に、ディセーブルになります。
- RHEL 6.0 アップデイト
- RedHawk 6.0 インストール
  - RHEL によって設定されたデフォルトファイアウォールを停止させます。
  - RHEL によって設定された SELinux を停止させます。
  - すべてのコンカレント RedHawk パッケージをインストールします。
  - 自動的にグラフィックスハードウェアを検出します。
  - NVIDIA と ATI の専用のグラフィックスドライバーは、同時にインストールすることは出来ません。
  - ネットワーク最新情報ユーティリティー（NUU）をインストールします。
  - 新しいオプションプロダクト： PCIe から VME へのブリッジ
- RedHawk 6.0 アップデイト（オプション）
- オプションプロダクトインストール（オプション）

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

RedHawk6.0 の変更点

- カーネル（スレッド）モジュールが増えました。
  - 5.4 → 178 カーネルモジュール。
  - 6.0 → 1857 カーネルモジュール。
- NUMA library 変更（v2 vs. v1）
  - numa_bind(nodemask_t* nodemask)
  - numa_bind(structbitmask* nodemask)
  - -DNUMA_VERSION1_COMPATIBILITY
- ksoftirqd ディフォルト優先度が変更（98 vs. 95）
  - CONFIG_SOFTIRQ_PRI（または 起動後）
- ワークキューデーモンのディフォルト優先度が変更(SCHED_RR/5 vs. SCHED_OTHER)。
  - CONFIG_WORK_PRI

RedHawk 6.0 の新しい機能

- Real-time にオプティマイズされたグラフィクス専用デバイスドライバ
  - NVIDIA グラフィクスドライバ 270.41 (vs. 190.42)
  - ATI グラフィクスドライバ 8.773 (ファーストリリース)
  - xorg.conf ファイルは自動的に生成されるようになりました。
- GPU コンピューティング・エンジン
  - CUDA 4.0 (vs. 3.0)は、標準提供
    - CUDA 4.0 は、64-bit 版のみサポートされます。
  - ATI Stream は、ご要求に応じて提供されます。 (RIQ)
- **RedHawk に特有な機能**
  - **mlockall_pid()**
    - 別のプロセスをメモリ常駐することができます。
    - ソースコードなしでリアルタイムの応答を改善することが出来ます。
    - 概要

      #include <mlockall_pid>

      int mlockall_pid(intflags, pid_tpid);

      int munlockall_pid(pid_tpid);

      gcc[options …] file –lccur_rt…
    - 例

      mlockall_pid(MCL_CURRENT|MCL_FUTURE, p1);

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

- ページフォールト発生時に SIGBUS シグナルで通知させることができます。
  - 概要

    CONFIG_PROC_PID_SIGBUS_PAGEFAULTS

    mlockall(MCL_CURRENT|MCL_FUTURE);
  - 例

    $ ./simulation &

    [1] 1092

    $ echo 1 > /proc/1092/sigbus_pagefaults
- RAM ディスクシールディング
  - 概要

    CONFIG_BLK_DEV_RAM_NUMA
  - 例

    $ echo 2 > /proc/driver/ram/ram1/nodemask ←NUMA node 1 (0x10)

    $ mke2fs –jvm0 /dev/ram1 4096

    $ mkdir/mnt/ramdisk

    $ mount –t ext3 /dev/ram1 /mnt/ramdisk
- NUMA シールディング
  - 概要

    どんなシールドコマンドの「CPULIST」でも、「n」接頭辞を使うことによって、

    すべての NUMA ノードでＣＰＵを指定することができます
  - 例

    shield -i0-2　　←CPU のディフォルト

    shield -a n1,n3

    shield -l c1 -p c1,c3,n2
- run コマンドの新しいスケジューリングポリシー
  - 対話型でないジョブのための SCHED_BATCH （P0）
  - 非常に低いプライオリティジョブのための SCHED_IDLE （P0）
  - 概要

    --policy|-s: fifo, rr, other, batch, idle
  - 例

    run -s idle ./statmon

    run –s batch –p 1257

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

- **新しい SCHED_RESET_ON_FORK 属性 のサポート**
  - ✧ チャイルドプロセスは、リアルタイムスケジューリングポリシーを継承せず、SCHED_OTHER で生成されます。
  - ✧ run コマンドの-r オプションでサポートされます 。
  - ✧ 概要

    --resetonfork|-r　但し--policy|-s=fifo|rr と同時に使用する事
  - ✧ 例

    run -r -s fifo -P 80 ./statmon

    run -r 5 ←間違った使用法（ポリシーが指定されていない）
- **NUMA CPU トポロジーを見る新しい run コマンドオプション**

$ run -M n

Node	MemSize	MemFree	Cpus
0	2048 MB	17 MB	0x0000,00000fff
2	2048 MB	219 MB	0x0000,00fff000
4	2048 MB	295 MB	0x000f,ff000000
6	2048 MB	190 MB	0xfff0,00000000

$ run -M c

Node	MemSize	MemFree	Cpulist
0	2048 MB	17 MB	0-11
2	2048 MB	219 MB	12-23
4	2048 MB	295 MB	24-35
6	2048 MB	190 MB	36-47

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

RedHawk のリアルタイム機能

CPU シールディング

コンカレント社は、ユーザーが選択した CPU 上で、割込みとシステムデーモンの引き起こす予想できない処理からシールドする方法を開発しました。

マルチコアシステムでは、任意の CPU に、リアルタイムプライオリティのタスクを割付け、他のＣＰＵに割り込みとシステムデーモンを割付けることによって、最も良いプロセスディスパッチレイテンシィを達成することができます。

この機能は、/proc ファイルシステムとコマンド shield(1)で、使うことが出来ます。

shield

3 種類の CPU 保護属性の設定/表示を行う

irqs	デバイス割り込み
ltmrs	local タイマー割り込み(APIC)
procs	無関係なプロセスの割付 (キャッシュのフラッシュやコンテキストスイッチ等)

cpu

システムの CPU の状態の設定/表示を行う

論理プロセッサの使用/非使用を行う

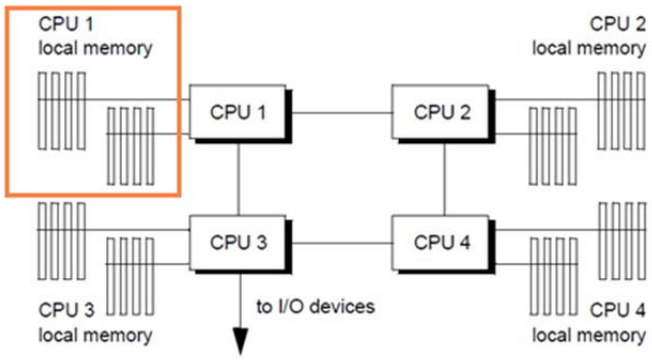
```
# shield -a3
# shield
     CPUID       all      irqs     ltmrs     procs
-------------------------------------------------
         0        no        no        no        no
         1        no        no        no        no
         2        no        no        no        no
         3       yes       yes       yes       yes
# cpu -d2
   log id
  (phys id)  state  shielding
  ---------  -----  -------------
    0 (0)    up     none
    1 (0)    up     none
    2 (1)    down   none
    3 (1)    up     proc irq ltmr
```

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

メモリシールディング

NUMA システム上では、特定の CPU に実装されているメモリ上に、クリティカルな高いプライオリティのタスクを割付け、他のプロセスの引き起こすメモリフォルト割込みやバストラフィックからリアルタイムプロセスをシールドすることが出来ます。

この機能は、コマンド shield(1)と run(1)で、使うことが出来ます。




```
# /usr/bin/shield -m1 -a 0-3 -c
CPUID   irqs    ltmrs   procs   mem
------------------------------------------------------------
    0   yes     yes     yes     yes
    1   yes     yes     yes     yes
    2   yes     yes     yes     yes
    3   yes     yes     yes     yes
    4   no      no      no      no
    5   no      no      no      no
    6   no      no      no      no
    7   no      no      no      no
    8   no      no      no      no
    9   no      no      no      no
   10   no      no      no      no
   11   no      no      no      no
   12   no      no      no      no
   13   no      no      no      no
   14   no      no      no      no
   15   no      no      no      no
```

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

プロセッサアフィニティ

複数プロセスがマルチコア上で実行されるリアルタイムアプリケーションでは、システム上のすべてのプロセスのＣＰＵ割付についての明示的なコントロールが必要です。

この機能は、ライブラリルーチン mapadvise(3)とコマンド run(1)で、使うことが出来ます。

run

プロセス（あるいはグループ）に対して以下の設定/表示を行う

（すでに実行中のプロセスでも可能）

プロセス優先度

スケジューリングポリシー

タイムクオンタム

実行可能 CPU の割り当て

実行 CPU の割り当て

nice 値の表示/設定

メモリロック

```
# run -q list
-20 => 300ms
-19 => 290ms
-18 => 280ms
-17 => 270ms
-16 => 270ms
  :      :
 -1 => 150ms
  0 => 150ms
  1 => 140ms
  2 => 130ms
  :      :
 18 =>   10ms
 19 =>   10ms
```

```
# run -i -u0
Pid    Bias   Actual  Policy  Pri  Nice  Name
1      0xb    0x3     other   0    0      init
2      0x1    0x1     fifo    99   0     migration/0
3      0x2    0x2     fifo    99   0     migration/1
4      0x0    0x0     fifo    99   0     migration/2
5      0x8    0x0     fifo    99   0     migration/3
6      0xb    0x3     fifo    1    0     keventd
7      0xb    0x3     fifo    98   0     kbottomd
8      0x1    0x1     fifo    98   19    ksoftirqd/0
9      0x2    0x2     fifo    98   19    ksoftirqd/1
10     0x0    0x0     fifo    98   19    ksoftirqd/2
11     0x8    0x0     fifo    98   19    ksoftirqd/3
12     0xb    0x3     other   0    0      kswapd
13     0xb    0x3     other   0    0      bdflush
14     0xb    0x3     other   0    0      kupdated
15     0xb    0x3     other   0    0      jfsIO
16     0xb    0x3     other   0    0      jfsCommit
17     0xb    0x3     other   0    0      jfsSync
18     0xb    0x3     other   0    0      pagebufd

20     0x2    0x2     other   0    0      pagebuf/1
21     0x0    0x0     other   0    0      pagebuf/2
22     0x8    0x0     other   0    0      pagebuf/3
23     0xb    0x3     other   0    0      khubd
:      :      :       :       :    :      :
```

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋２丁目 19 番６号　柳橋ファーストビル４階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

concurrent

ユーザーレベルプリエンプション制御

マルチコア上で、共有されたデータを保護するために、最も効率的なメカニズムはスピンロックです。

しかし、もしアプリケーションが、スピンロックを保持している間に、プリエンプションされる可能性があるならば、スピンロックは、アプリケーションによって効率的に使うことが出来ません。 効果的なままでいるために、 RedHawk はアプリケーションが高速に、プリエンプションを禁止/許可することをメカニズムを提供します。

高速ブロック / ウェイクサービス

多くのリアルタイムアプリケーションは、複数の協調プロセスで構成されています。 これらのアプリケーションは、プロセス間の同期化をするための効率的な手段を必要とします。

コンカレント社は、高速なブロック / ウェイクサービス postwait(2) と server_block(2) を開発しました。

RCIM サポート

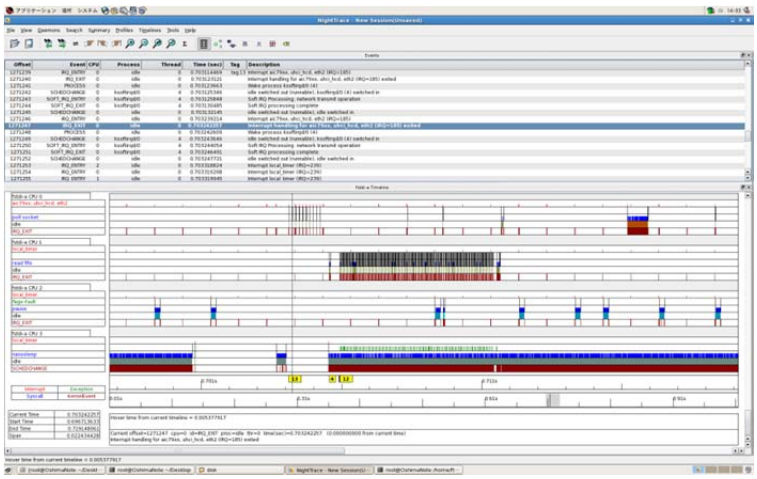
RedHawk は、外部イベントに瞬時に反応する必要のあるタイムクリティカルアプリケーションのために設計された多機能カード「リアルタイム・クロック&インタラプト・モジュール（RCIM: Real-Time Clock & Interrupt Module)」をサポートします。 8 個のプログラム可能なタイマーと 12 個の入出力外部割込みラインを利用できます。マルチシステムアプリケーションを同期させるために他の iHawk に対して割り込みソースを分配することが可能です。 RCIM には、複数のシステム全体に同じ時間を提供する高分解能同期クロックが含まれています。 RCIM 上のタイムスタンプは、RedHawk が高分解能 NTP を同期させるために供給することが可能です。 RCIM オプションには、外部からの入力無しに正確な時間を保持できる GPS 標準時間と高安定性水晶発振器が一体となり、同期させるための GPS モジュールが含まれます。

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

カーネルトレースファシリティ

RedHawk には、カーネルトレースを行う機能が標準で組み込まれています。カーネルトレース機能は、カーネルを通過する際、トレースポイントデータをメモリバッファに更新し、その情報を NightTrace ツールが読み出します。

このツールはリアルタイムに動作し、ユーザプログラムに変更を行う必要はありません。



ptrace 拡張

Linux 上の ptrace のデバッギングインタフェースは NightView デバッガの性能をサポートするために拡張されました。

- デバッガプロセスによって現在停止状態にあるプロセスがメモリを read/write 出来ない機能
- デバッガによってデバッグされているプロセスのシグナルの下位グループだけを追跡する機能
- デバッガによってデバッグされているプロセスが効率的に実行を再開する機能
- デバッガプロセスが生成するすべてのチャイルドプロセスを自動的にデバッガで制御する機能

カーネルプリエンプション

高いプライオリティプロセスがカーネルの中に現在実行しているより低いプライオリティプロセスをプリエンプションする機能が提供されています。

スタンダード Linux では、プリエンプションが行われないため、より長いワーストケースプロセスディスパッチレイテンシィを発生させます。

リアルタイムスケジューラ

いくつのプロセスがシステムでアクティブであっても、リアルタイムスケジューラは一定時間のコンテキストスイッチ時間を提供します。

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

ローレイテンシィ拡張

カーネルによって使用されている共有データ構造を守るために、カーネルはスピンロックと セマフォで、これらの共有データ構造にアクセスするコードパスを保護します。

スピンロックでロックすることはプリエンプションと割り込みを行うために必要です。

スピンロックが保持される間は保護されたコードパスは、実行不能でなくてはなりません。ワーストケースレイテンシィのプリエンプションの低減の研究が行われ、ローレイテンシィ拡張は、もっと良い割り込み応答時間を提供するためにシナリオからアルゴリズムで識別されるワーストケースのプリエンプションを修正します。

プライオリティインヘリタンス

Sleeping-Wait を排他制御メカニズムとして使用したセマフォはプライオリティ逆転の問題を発生させました。

クリティカルセクションを実行している低いプライオリティのプロセスが、高い優先順位プロセスの実行を妨げるとき、プライオリティ逆転が発生します。

プライオリティインヘリタンスは、クリティカルセクションを実行している低いプライオリティのプロセスに、一時的に高優先度待機プロセスの優先度を与えることで回避され、クリティカルセクションを去るまで、クリティカルセクションで実行しているプロセスが実行を続けるために十分なプライオリティを持っていることを保証します。

高解像度プロセスアカウンティング

標準的な Linux カーネルでは、システムは非常に粗い粒度のメカニズムを使ってプロセスのＣＰＵ実行時間を報告します。 これは特定のプロセスに加算されたＣＰＵ時間の量が非常に不正確であり得ることを意味します。 高精細プロセスのアカウントファシリティは、アプリケーションの最良のパフォーマンスモニタリングを行うことができ、非常に正確なＣＰＵ実行時間アカウントメカニズムを提供します。 このファシリティは、コンカレント社によって供給されるデバッグカーネルとトレースカーネルで監視する標準的な Linux ＣＰＵアカウントによって使われます。

ケイパビリティサポート

Pluggable Authentication Module (PAM) はユーザーに特権を割り当てて、そして認証プログラムを再コンパイルすることを必要としないで認証ポリシィを設定するためのメカニズムを提供します。

このスキームの下で、非 root のユーザーが特権を必要とするアプリケーションを走らせるように設定されることができます、ただし root は常に特権を許されます。

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

concurrent

例えば、ページをメモリにロックする機能は個別のユーザーあるいはグループに割り当てられることができる１つの前もって定められたプリビレッジによって提供されます。
特権はコンフィギュレーション・ファイルで定義された role を通して与えられます。

カーネルデバッガ

オープンソースカーネルデバッガ kdb は RedHawk Linux のデバッグカーネルでサポートされています。

カーネルコアダンプ / クラッシュ分析

kexec と kdump オープンソースのパッチは、別のカーネルにロードして、そしてクラッシュダンプを取り込むことができます。
そしてクラッシュユーティリティーがダンプを分析するために提供されています。

ユーザーレベルスピンロック

RedHawk Linux のビジィ待ちの排他制御ツールは、低いオーバーヘッドのビジィ待ちの排他制御変数（スピンロック）とスピンロックを初期化して、ロックして、アンロックして、そして問い合わせることを可能にするマクロのセットです。
ユーザーレベルスピンロックは、ユーザーレベルプリエンプション制御で使われなくてはなりません。

Hyper-threading

ハイパースレッディングは、インテルペンティアム Xeon プロセッサの機能です。 それは一つの物理的なプロセッサが２つの論理的なプロセッサとしてオペレーティング・システムに認識されます。
結果的に、それぞれのチップはデュアルＣＰＵの SMP であるように、２つのプログラムが、それぞれのＣＰＵチップの中で同時に走ります。
RedHawk Linux は、ハイパースレッディングへのサポートを含んでいます。

XFS ジャーナリング・ファイルシステム

SGI からの XFS ジャーナリング・ファイルシステムが RedHawk Linux に実装されています。 ジャーナリング・ファイルシステムはジャーナル（ログ）をトランザクション記録のために使います。 システムクラッシュの場合、バックグラウンドプロセスはリブート時にジャーナルからファイルシステムまで更新をコピーして終わります。
これは劇的に、回復時間を減らして、ファイルシステムチェックの複雑さをカットします。

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋２丁目 19 番６号　柳橋ファーストビル４階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp

concurrent

POSIX リアルタイム拡張

RedHawk Linux は、ISO/IEC 9945-1 で定義された、 POSIX リアルタイム拡張インタフェースの大部分をサポートします。

- ユーザ優先度制御
- プロセスメモリロック
- メモリマップファイル
- 共有メモリ
- メッセージキュー
- カウンティングセマフォ
- リアルタイムシグナル
- 非同期入出力(asynchronous I/O)
- 同期入出力(synchronized I/O)
- 高解像度タイマー

詳細は、RedHawk User Guide をご覧下さい。

RedHawk 6.0 プロダクト

RedHawk6.0 には、32 ビット版と 64 ビット版があり、以下のプロダクトが御座います。
詳細は、コンカレント日本株式会社まで、お問い合わせ下さい。

Product	RHEL6	RedHawk6	NightStar	Architect	Computer	RCIM
iHawk	○	○	Option	×	○	○
RedHawk Server	○	○	○	○	×	×
RedHawk Embedded	○	○	○	○	×	×

コンカレント日本株式会社
〒111-0052　東京都台東区柳橋 2 丁目 19 番 6 号　柳橋ファーストビル 4 階
TEL：03-3864-5713 FAX 03-3864-0898
製品に関するお問い合わせ info@ccur.co.jp