

これにより、コンピュータの操作性が格段に向上した。1970年代から1980年代にかけて、ほとんどのコンピューターにはVDT技術と入力用の電子キーボードが搭載されていた。その後、VDTに代わってCRTやLCDが採用されるようになり、電子キーボードも汎用コンピュータの標準となった。

今日、私たちはコンピュータを使うたびにキーボードを使っています。キーボードのレイアウトのほとんどはタイプライターのものが残っており、使い方も同じです。しかし、電子機器の新しい時代のおかげで、キーボードにはさまざまな形があります。一般的なプラスチック製のキーボード、折り畳み式のキーボード、バックライト付きのキーボード、さらにはレーザーキーボードまで。

キーボードレイアウト

一般的なキーボードレイアウトは、QWERTYという文字が最初の5文字であることから、QWERTYキーボードと呼ばれています。QWERTY配列は、タイプライターの時代に、初期のタイプライターの機械的な限界から、タイピストのタイピング速度を遅くするために意図的に設計されたものです。これは主に、各キープレスの間の時間を短縮し、プリントヘッドに十分な時間を与えてジャムを起こさないようにするためである。

QWERTY配列は、現在もすべてのキーボードに採用されています。

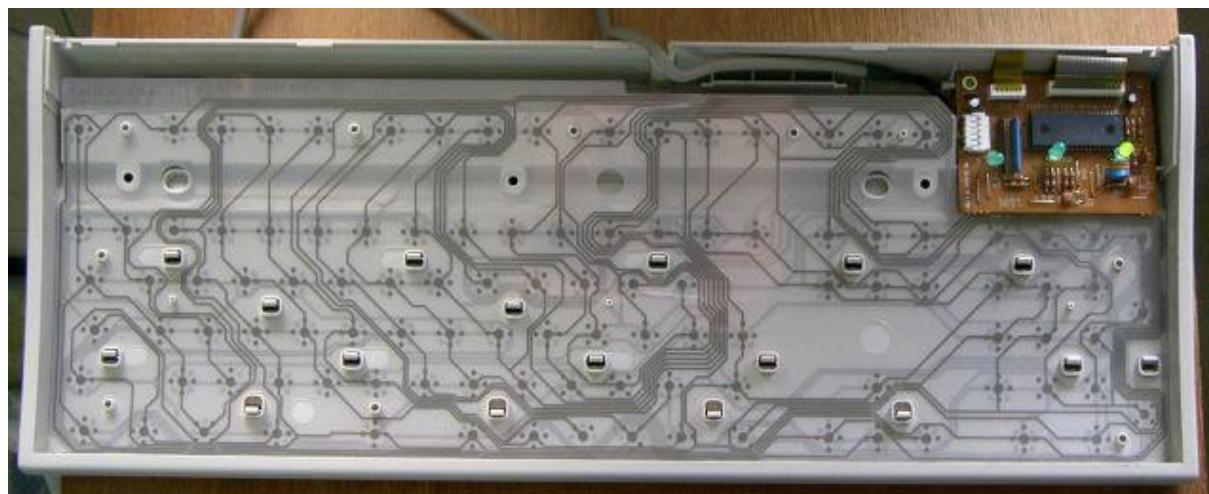
キーボード内部

キーボードのキーを押すと、実際に何が起こるのでしょうか？キーボードは、どのキーが押されているかをどうやってプログラムに伝えるのでしょうか？今、読まれている文章は、キーボードで入力されたものです。キーボードはどのようにしてこれを行うのでしょうか？見てみましょう。

注：具体的な内容は、キーボードの種類やモデルによって異なります。そのため、ここでは一般的な102キーのキーボードのみを取り上げています。

ケースを開く

キーボードが複雑なプリント基板から、マイクロプロセッサを搭載した一体型の基板になったことに驚かれるかもしれません。キーボードを開いてみると、このようになっています。



そう、それです。いかにシンプルであるかがわかります。1枚の回路基板とグリッドです。上の写真ではグリッドが少し見づらいかもしれません。しかし、よく見ると、グリッドのポイントが見え、そのポイントが一般的なキーボードのキーポジションと一致していることに気づくかもしれません。これを「キーマトリクス」と呼びます。ほとんどのキーボードでは、キーマトリクスを構成する回路がグリッドの各ポイントの間で途切れています。キーマトリクスのあるポイントの上にキーがあることを知って、キーを押すと、そのポイントにあるスイッチが押されて横の回路が完成し、電流が流れるようになります。キーの機械的な動きによって生じる線の振動はバウンスと呼ばれ、キーボードエンコーダーとして知られるキーボード自身のマイクロプロセッサーによってフィルタリングされます。少し複雑に感じられるかもしれませんが、ご安心ください。次の2つのセクションでは、すべてをより詳しく見ていきます。

キーボードエンコーダ

キーボードに搭載されているマイクロプロセッサーは、インテル社の最初のマイクロコントローラーでもある、オリジナルのインテル8048が使用されている。このコントローラーはキーボード・エンコーダーと呼ばれています。キーボードエンコーダーの種類は、キーボードによって大きく異なります。キーボード・エンコーダーには何百種類もの種類がありますが、基本的にはどれも同じことをしています。

キーグリッド内の行と列は、キーボードエンコーダーの8ビットI/Oポートに接続されています。キーが押されると、キーグリッド内のその場所にあるスイッチが閉じ、電流が流れ回路が完成します。この電流は、キーの位置に対応するポートのキーボードエンコーダのピンを有効にします。このように、コントローラはポートをスキャンするだけで、キーが押されているかどうかをポートラインがアクティブかどうかで確認することができます。

キーが押されていると、キーボードのエンコーダはROM (Read Only Memory) の文字マップからその文字のスキャンコードを調べ、内部の16バイトのメモリに保存する。キーボードのプロセッサには、独自のタイマーと33の命令セットがあり、128Kの外部メモリにアクセスすることもできる。タイマーを使って、キーが押されたかどうかを、ユーザーの入力かバウンスかで判断します。バウンドした場合は、通常、人間が入力できるよりもはるかに速い速度になります。タイマーが0になつてもキーが押されたままであれば、リセットされ、文字が内部の16バイトのバッファに挿入される。

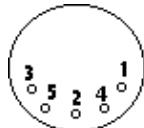
ここで重要なのは、通信可能なキーボードコントローラーが2つあるということです。キーボードの内部にあるキーボードエンコーダと、マザーボード上にあるキーボードコントローラです。もう1つのコントローラについては後ほどご紹介します。）今のところ、2つのコントローラーがあり、キーボードエンコーダーはそのうちの1つであることを覚えておいてください。

キーボードエンコーダは、キーボードプロトコルで定義された方法でシステムと通信します。その内容を見てみましょう。

キーボードプロトコル

キーボードエンコーダは、データをバイトとしてマザーボードのオンボードキーボードコントローラに送信します。データの送信方法は、キーボードのインターフェースで使用されているプロトコルによって異なります。これは通常、5ピンDINコネクタ、6ピンMini-DINコネクタ、USBコネクタ、SDLコネクタ、または赤外線（IR）インターフェースを使用したワイヤレスです。

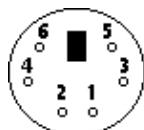
AT/XTキーボードに使用する5ピンのDINコネクタは、通常、コンピュータの背面にあり、次のような形をしています。



1: クロック 2: データ 3: 該当なし 4: グラウンド 5: Vcc (+5V)

マザーボードは、Vcc端子とGround端子を介してPSU (Power Supply Unit) から電源を供給します。クロック端子は、キーボードのデータとシステムクロックとの同期に使用されます。キーボードからのデータは、データピンを介してシリアルデータとして送信されます。

PS/2キーボードに使用されている6ピンのMini-DINコネクタは非常によく似ています。



1: データ 2: N/A 3: グラウンド 4: Vcc (+5V) 5: クロック 6: N/A

ここでは特に新しいことはありません。DINは特に何かを表しているわけではなく、これを開発した標準化団体 (Deutsches Institut für Normung、英語ではGerman Institute for Standardization) を意味しています。

SDL (Shielded Data Link) コネクタは非常によく似ています。



A: N/A B: データ C: グラウンド D: クロック E: Vcc (+5V) F: N/A

ユニバーサル・シリアル・バス（USB）コネクターは、さまざまな機器で使用されている規格です。USBデバイスを直接扱うことは、かなり複雑なテーマです。USBデバイスには4つのピンしかありません。1: Vcc (+5V)、2: Data-、3: Data+、4: Ground。

USBレガシーサポートは、USBポートを備えた最近のほとんどのコンピュータで採用されています。これは、これらのコンピュータのマザーボードが、USBキーボードやマウスをPS/2キーボードやマウスとしてエミュレートできることを意味します。このため。PS/2互換のインターフェースを使用したUSBキーボードやマウスとの通信は機能します。言い換えれば、私たちのようにUSBキーボードやマウスを持っていても心配する必要はありません。このチュートリアルのコードとデモは、マザーボードが提供するエミュレーションおかげで問題なく動作します。

ご覧のように、キーボードとコンピュータの間のインターフェースはそれほど複雑ではありません。キーボードコントローラとキーボードエンコーダの間で、データをビットとして送信する方法を提供しているだけです。そのデータは

マザーボード上のオンボードまたは一体型のキーボードコントローラーにルーティングされます。キーボードコントローラーが制御します。

キーボードコントローラー

システムケース内で使用されているキーボードコントローラーは、通常、オリジナルの8042キーボードコントローラーの形をしています。キーボードコントローラーは、キーボードのプロトコルを介してキーボードエンコーダーとのインターフェースを提供します。ほとんどの新しいシステムでは、キーボードコントローラーは独立したICではなく、フロッピーディスクコントローラー (FDC)、パラレルポートインターフェース、シリアルポートインターフェース、マウスインターフェースを含むマザーボードのスーパー入出力 (IO) コントローラーの一部となっています。最近のシステムでは、スーパーIOコントローラーは、マザーボードのサウスブリッジにあるISA (Industry Standard Architecture) ではなく、LPC (Low Pin Count) バスを使用しています。

Scan Codes

スキャンコードとは、キーの状態を表すデータパケットのこと。キーが押されたり、離されたり、押し続けられたりすると、スキャンコードがコンピューターのオンボードキーボードコントローラーに送られる。スキャンコードには2種類あります。メイクコードとブレークコードである。メイクコードはキーが押されたときに送信され、ブレークコードはキーが離されたときに送信されます。メイクコードとブレークコードは、キーボードの各キーに固有のものです。すべてのスキャンコードを表す数字のセットがキーボードのスキャンコードセットです。

キーボードが使用できるスキャンセットは通常3種類あります。しかし、スキャン値はランダムなので、どのスキャンセットを使用しているかを簡単に判断する方法はありません。そのため、ルックアップテーブルを使って、スキャンコードが示すキーを決定する必要があります。

それでは、スキャンコード表を見てみましょう。注意：これらの表は重要です。これらのテーブルは重要です！キーボードのどのキーが押されたかを判断するのに必要です。また、これらの表に記載されているスキャンコードはすべて16進法です。

これらの表はかなり大きいので、別のリソースとして置くことにしました。表はこちらのリソースセクションでご覧ください。

例を挙げてみましょう。キーボードのshift+Aキーを押すと、コンピューターに送られるメイクコードはどうなるでしょうか？これを理解するために、一連の流れを見てみましょう。まず、シフトキーが押され、次にAキーが押されます。次にAキーを離し、続いてシフトキーを離します。スキャンコードセットが最近のキーボードのデフォルトのスキャンコードセットであると仮定すると、左シフトキーのメイクコードは0x12、ブレークコードは0xF0と0x12です。また、Aキーのメイクコードは0x1C、ブレークコードは0xF0と0x1Cです。したがって、このイベントが発生すると、以下のスキャンコードがコンピュータに送信されます。

Key events:	shift down	A down	A released	Shift released
Scan codes:	0x12	0x1C	0xF0	0x1C
			0xF0	0x12

上記を見ると、送信されるスキャンコードは、0x12、0x1C、0xF0、0x1C、0xF0、0x12となることがわかります。

キーを押したままにしておくと、そのキーはタイプコードになります。つまり、キーを離すか別のキーを押すまで、キーボードはキーのコードを送り続けるのです。試してみてください。お気に入りのテキストエディターを開き、あるキーを押し続けます。しばらくすると、同じ文字が現れ、その後、その文字が連続して表示されます。タイプディレイは、タイプモードに入るまでの待ち時間を決定し、タイプレートは、コンピュータに送信する1秒間の文字作成コードの量を決定します。タイプマッチングモードでは、文字データはバッファリングされません。複数のキーを押し続けた場合、最後に押したキーだけがタイプマッチングになります。

スキャンコードは、私たちにとって非常に重要なものです。スキャンコードがオンボードのキーボードコントローラーに送信されると、キーボードコントローラーはスキャンコードを内部メモリーに格納します。その後、キーボードコントローラーは、割り込み要求 (IRQ) ラインをハイに切り替えます。割り込みラインがPIC (Programmable Interrupt Controller) によってマスクされていない場合は、これによってIRQ 1が発生します。IRQがマスクされていても、リードバッファはソフトウェアで読み取るので、スキャンコードを読み取って、どのキーが離されたのか、押されたのかを判断すること

キーボードインターフェース。デバイスドライバの開発

この章では、すでに多くのことを取り上げてきました。インターフェース機器としてのキーボードの歴史、QWERTYキーボードのレイアウト、キーボードの内部を見て、その仕組みと主要部品を確認しました。また、スキャンコードセットやキーボードのプロトコルについても見てきました。まだすべてを理解していないなくても心配いりません。次のいくつかのセクションでさらに詳しく説明します。また、キーボードのデバイスドライバーも開発する予定です。すごいでしょう？このセクションのすべてのコードは、最終的なデモにも使用されます。

キーボード・インターフェーシング ポーリング

前のセクションで、キーボードを操作する際には2つのコントローラがあることを覚えていますか？つまり、キーボード内部のキーボードエンコーダーと、マザーボード上のキーボードコントローラです。この章では、1つのハードウェアデバイスを制御するために、複数の異なるコントローラとのインターフェースが必要になります。そうですね。これらのコントローラの両方と通信することができるので、まあ、そんな感じです。キーボードエンコーダーにコマンドを送信すると、オンボードのキーボードコントローラにもコマンドが送信されますが、キーボードプロトコルを介してキーボードエンコーダーに再ルーティングされます。

よし、これで両方のコントローラと通信できるぞ。これは楽しいですね。両方のコントローラが相互に機能しているということは、相互に通信もしているということです。キーボードエンコーダーは、さまざまなコードをオンボードのキーボードコントローラに送信して保存します。これはスキャancodeであったり、エラーコードであったりします。これにより、キーボード・エンコーダーとオンボード・コントローラの両方から情報を受け取ることができます。

これらの通信は、IOアドレス空間にマッピングされたコントローラのポートに対して、IN命令とOUT命令を使って読み書きするだけで行われます。これらのポートが何であるかを気にする必要はありませんでしたが、IOマッピングがコントローラでどのように機能するかを理解することは、ここでより重要になります。

これは、キーボードとのインターフェースのひとつです。キーが押されているか、押されていないかなどを確認するために、コントローラと手動で通信することができます。これを「キーボードのポーリング」と呼びます。このようにして、キーボードコントローラにポーリングすることで、キーボードから最後のスキャancodeを取得することができます。

キーボードのインターフェイス。割り込み要求(IRQ)

PICのチュートリアルで、キーボードコントローラが割込み線を使用するように設定できることを覚えていますか？キーが押されたり離されたりしたときに、キーボードコントローラがIRQ 1を発行するように設定することができます。これは、キーボードとの最も一般的なインターフェース方法です。

IRQ1が起動したときは、必ずキーボードコントローラにスキャancodeが実際に送られたかどうかをテストする必要があります。これは、キーボードコントローラをポーリングして最後のスキャancodeを取得することで行います。

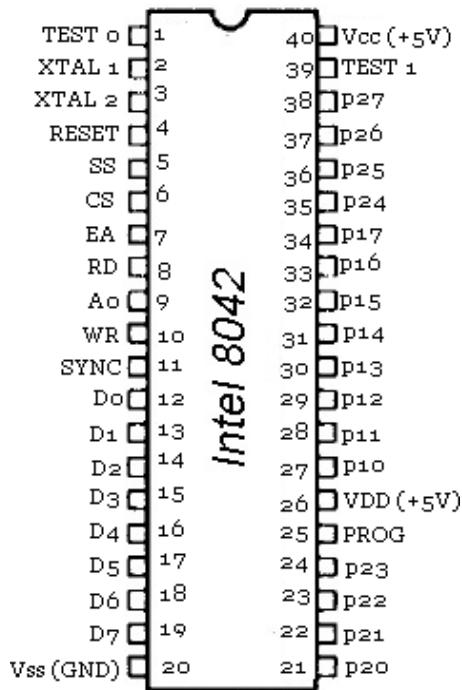
詳細 8042キーボードマイクロコントローラ



オリジナルの8042マイクロコントローラ

キーボードエンコーダーとのインターフェースを担うマイクロコントローラです。キーボードコントローラは、8042マイクロコントローラから始まったマイクロコントローラファミリーの一部です。最近のコンピュータでは、キーボードコントローラは独立した集積回路（IC）ではなく、その機能はマザーボード自体にエミュレートされています。つまり、コントローラの機能は、マザーボードのチップセットに組み込まれているのです。

キーボードコントローラは、2つのモードで動作します。ATコンパティブルモード」と「PS/2コンパティブルモード」です。コントローラがどちらのモードで動作するかによって、外界とのインターフェース方法が異なります。まず、コントローラを見てみましょう。



うん。その通りです。P10-P17ピンがコントローラーの入力ポート。ピンP20-P27は、コントローラの出力ポートです。そして、ピンT0-T1はコントローラのテストポートです。これらのピンの正確な意味は、コントローラの動作モードによって異なります。

これらのポートを操作するためのコマンドがありますので、後で詳しく見てみましょう。

他のピンのほとんどは、私たちにとって重要ではありません。ここに追加したのは、あくまでも補足のためであり、あなたがそれらを知る必要はありません。

XTAL 1とXTAL 2は、水晶振動子の入力端子です。CLKが外部から駆動されている場合、XTAL 1はグランドに接続することもできます。同様に、CLKが外部から駆動されている場合、XTAL 2はCLKに接続することができます。

RESETはLow(0)になるとコントローラをリセットします。

SSは、マイクロコントローラのシングルステップ端子です。CSは、データレジスタポートのインターフェースに使用されるチップセレクト端子です。EA (No, not the company;) は外部アクセス入力端子です。OTP (ワンタイム・プログラマブル) ROMをディセーブルにして、外部からコントローラにコマンドを送れるようにします。

RD 出力イネーブル入力：データレジスタポートのインターフェイスに使用されます。A0はコマンド/データ・レジスタ・セレクト入力です。データ・レジスタ・ポートのインターフェーシングに使用します。WRはライトイネーブル入力ラインで、データレジスタポートのインターフェーシングに使用されます。

SYNCはクロック出力信号です。D0～D7はデータレジスタポートのインターフェイスに使用します。GNDはグランド端子(Vss)です。Vddは、+5V入力端子である。PROGは、I/Oエクスパンダアクセス時に8243へのアドレス/データストローブとして使用される。Vccはもう一つの+5V入力端子である。

キーボードコントローラは、キーボードの動作を制御するためのインターフェースを提供します。これは、ポートI/O空間にマッピングされたポートを介して、キーボードコントローラと通信することで行います。ご存知のように、キーボードコントローラと通信するためには、INとOUTの命令を使い、そのマッピング方法を知る必要があります。では、見てみましょう。

ポートマッピング

i86アーキテクチャでは、キーボードとの通信に以下のポートが使用されています。

Keyboard Controller Ports		
Port	Read/Write	Description
Keyboard Encoder		

0x60	Read	Read Input Buffer
0x60	Write	Send Command
Onboard Keyboard Controller		
0x64	Read	Status Register
0x64	Write	Send Command

このテーブルは悪くないと思いますよ;) 基本的にはキーボードエンコーダにコマンドを送るには、ポート0x60にコマンドバイトを書き込みます。しかし、これを行う前に、安全のためにキーボードコントローラのステータスレジスタのビット0（出力バッファフル）が0であることを確認する必要があります。もし、キーボードコントローラのステータスレジスタのビット1（入力バッファフル）が1であれば、データは入力バッファに入っていて、読める状態になっています。ポート0x60から読み出すと、キーボードエンコーダからこのデータを取得することができます。キーボードエンコーダから読み込んだデータは、通常はキーボードから得られます。しかし、マイクロコントローラを再プログラムして、特定の値を返すようにすることもできます。

ポート0x64に値を書き込むことで、オンボードのキーボードコントローラにコマンドバイトを送信することができます。

ポート0x64から読み出すと、キーボードコントローラのステータスバイトを取得することができます。

これらのことを行っていれば、これらのコントローラとの間でコマンドバイトやデータを読み書きするルーチンを簡単に提供することができます。ここでは、これらのコントローラが使用するIOポートを抽象化しています。

```
KYBRD_ENC_INPUT_BUF      =      0x60,
KYBRD_ENC_CMD_REG        =      0x60
```

```
KYBRD_CTRL_STATS_REG    =      0x64,
KYBRD_CTRL_CMD_REG      =      0x64
```

ここでは、コントローラのコマンドに関する知識を必要とするため、これらのコントローラを操作するためのルーチンについては、まだ説明しません。

レジスター

ステータスレジスター

これは、20番目のアドレスラインを有効にすることを説明したときに見覚えがあるかもしれません。ステータスレジスターを読み出すには、I/Oポート0x64から読み出すだけです。返される値は、特定のフォーマットに従った8ビットの値です。このフォーマットは、コントローラのモードに応じて少しずつ異なります。

ここにもう一度書いておきます。重要なものは太字にしました。

ビット0：出力バッファステータス

- 0: 出力バッファが空、まだ読まないでください
- 1: 出力バッファがいっぱい、読んでください
- い:)

ビット1：入力バッファステータス

- 0: 入力バッファが空、書き込み可能 1: 入力バッファが満杯、まだ書き込まないでください
- さい

ビット2：システムフラグ

- 0: パワーオンリセット後に設定
- 1: キーボードコントローラのセルフテスト (Basic Assurance Test, BAT) が正常に終了したときに設定されます。

ビット3：コマンドデータ

- 0: 入力バッファへの最後の書き込みはデータ (ポート0x60経由)
- 1: 入力バッファへの最後の書き込みがコマンド (ポート0x64経由) だった場合
- Bit 4: キーボードロック
- 0: ロックされて
- いる

1: ロックされていない

ビット5: Auxiliary Output

buffer full PS/2 Systems:

0: ポート0x60からの読み出しが有効かどうかを判断 有効な場合、0=キーボー

ドデータ 1: マウスデータ、ポート0x60からの読み出しが可能な場合のみ

ATシステム。

0: OKフラグ

1: キーボードコントローラからキーボードへの送信がタイムアウトした。これは、キーボードが存在しないことを示している可能性があります。

ビット6: タイムアウト

0: OKフラグ

1: タイムアウ
ト PS/2

一般的なタイムアウト

◦ AT:

キーボードからキーボードコントローラへの送信のタイムアウト。パリティエラーの可能性あり（この場合、ビット6と7の両方がセットされます）

ビット7：パリティエラー

0：OKフラグ、エラーなし

1：最終バイトのパリティエラー

キーボードの現在の状態を判断し、何ができる何ができないのかを知るために、ステータスレジスタを読み取る必要があります。例えば、キーボードが接続されていない状態で、キーボードにコマンドを送信したくはありません。そこで、コマンドを送信する前に、現在の状態を読み込んでテストしたいのです。

また、キーボードコントローラが反応するよりもプロセッサが命令を実行する方がはるかに速いということも考慮する必要があります。このため、キーボードコントローラが次のコマンドを実行できるようになるまで、待たなければならないことがあります。これを確認するには、ステータスレジスタを読み込んで、ビット0（Output Buffer Full）をテストし、次のコマンドを送信してもよいかどうかを確認する必要があります。これを行わないと、前のコマンドは破棄され、新しいコマンドが実行を開始しますが、これは好ましくないことです。

別のコマンドを送信したり、データを読み出す前に、コントローラの準備が整うのを待つことが重要です。

ステータスレジスターの読み書きには、ビットマスクを使うことができます。この章の最後に掲載したデモのものです。各ビットが上のリストの正しいビットとどのように一致しているかに注目してください。

KYBRD_CTRL_STATS_MASK_OUT_BUF	=	1,	//00000001
KYBRD_CTRL_STATS_MASK_IN_BUF	=	2,	//00000010
KYBRD_CTRL_STATS_MASK_SYSTEM	=	4,	//00000100
KYBRD_CTRL_STATS_MASK_CMD_DATA	=	8,	//00001000
KYBRD_CTRL_STATS_MASK_LOCKED	=	0x10,	//00010000
KYBRD_CTRL_STATS_MASK_AUX_BUF	=	0x20,	//00100000
KYBRD_CTRL_STATS_MASK_TIMEOUT	=	0x40,	//01000000
KYBRD_CTRL_STATS_MASK_PARITY	=	0x80	//10000000

すばらしいですね。というわけで、あとはポート0x64にあるキーボードコントローラのステータスレジスタを読み込めばいいわけです。そして、上のビットマスクに基づいて、好きなビットをテストして、そのステータスをチェックします。

つまり、キーボードコントローラのステータスレジスタから読み出すために必要なのは

```
///! read status from keyboard controller
uint8_t kybrd_ctrl_read_status () {
    return inportb (KYBRD_CTRL_STATS_REG);
}
```

読み取りと書き込み入力バッファ

コマンドを送信するには、まず、キーボードコントローラがコマンドを受信できる状態になっていることを確認します。これは、入力バッファが一杯になっているかどうかを確認することで行われます。これは、キーボードコントローラのステータスレジスタを読み、ビットをテストすることで行います。このビットが0であれば、バッファは空なので、コマンドバイトを送信します。（これらの情報はすべて、上に示したステータスレジスタのビットレイアウトの中にあることを覚えておいてください）

```
///! send command byte to keyboard controller
void kybrd_ctrl_send_cmd (uint8_t cmd) {
    //! wait for kkybrd controller input buffer to be clear
    while (1)
        if ( (kybrd_ctrl_read_status () & KYBRD_CTRL_STATS_MASK_IN_BUF) == 0 )
            break;

    outportb (KYBRD_CTRL_CMD_REG, cmd);
```

キーボードエンコーダーは、以下のように非常によく似ています。キーボードエンコーダーに送られたコマンドは、まずキーボードコントローラに送られることを覚えておいてください。このため、キーボードコントローラ自体がコマンドを受信できる状態になっていることを確認する必要があります。

```

//! read keyboard encoder buffer
uint8_t kybrd_enc_read_buf () {
    return inportb (KYBRD_ENC_INPUT_BUF);
}

//! send command byte to keyboard encoder
void kybrd_enc_send_cmd (uint8_t cmd) {
    //! wait for kkybrd controller input buffer to be clear
    while (1)
        if ( (kybrd_ctrl_read_status () & KYBRD_CTRL_STATS_MASK_IN_BUF) == 0)
            break;

    //! send command byte to kybrd encoder
    outportb (KYBRD_ENC_CMD_REG, cmd);
}

```

Keyboard Encoder Commands

ポート0x60にコマンドバイトを書き込むと、キーボードコントローラはその値をキーボードエンコーダに直接送信します。以下に、コマンドバイトの一覧を示します。

Command Listing	
Command	Description
0xED	Set LEDs
0xEE	Echo command. Returns 0xEE to port 0x60 as a diagnostic test
0xF0	Set alternate scan code set
0xF2	Send 2 byte keyboard ID code as the next two bytes to be read from port 0x60
0xF3	Set autorepeat delay and repeat rate
0xF4	Enable keyboard
0xF5	Reset to power on condition and wait for enable command
0xF6	Reset to power on condition and begin scanning keyboard
0xF7	Set all keys to autorepeat (PS/2 only)
0xF8	Set all keys to send make code and break code (PS/2 only)
0xF9	Set all keys to generate only make codes
0xFA	Set all keys to autorepeat and generate make/break codes
0xFB	Set a single key to autorepeat
0xFC	Set a single key to generate make and break codes
0xFD	Set a single key to generate only break codes
0xFE	Resend last result
0xFF	Reset keyboard to power on state and start self test

小さなコマンドはすべて上の表に記載されています。ここでは、より複雑なコマンドについて詳しく見ていきましょう。

Command 0xED - Set Light Emetting Diods (LED's)

このコマンドは、キーボードのLEDを設定するために使用します。ポート0x60に書き込まれた次のバイトでキーボードのLEDが更新され、以下のようなフォーマットになります。

ビット0：スクロールロックLED（0：オフ
1：オン） ビット1：ナムロックLED
(0：オフ 1：オン) ビット2：キャップ
スロックLED (0：オフ 1：オン)

その他のビットはすべて0でなければなりません。

このコマンドは、遊んでみると面白いですよ;) 以下は、デモがキーボードのライトを更新するために使用するルーチンの例です。パラメーターが真か偽かに基づいて、ビットをどのように設定またはクリアするかに注目してください。また、最初に

コマンドバイトをキーボードエンコーダに書き込み、次にデータバイトを書き込んでいることにも注目してください。これらは両方ともキーボード・エンコーダーのコマンド・レジスタに送られます。KYBRD_ENC_CMD_SET_LEDは、0xED（使用するコマンド・バイト）の定数です。魔法は使いません。）

```

//! sets leds
void kkybrd_set_leds (bool num, bool caps, bool scroll)

    { uint8_t data = 0;

        //! set or clear the bit
        data = (scroll) ? (data | 1) : (data & 1);
        data = (num) ? (num | 2) : (num & 2);
        data = (caps) ? (num | 4) : (num & 4);

        //! send the command -- update keyboard Light Emetting Diods (LEDs)
        kybrd_enc_send_cmd (KYBRD_ENC_CMD_SET_LED);
        kybrd_enc_send_cmd (data);
    }
}

```

Command 0xF0 - Set alternataate scan code set (PS/2 Only)

このコマンドは、使用するスキャンコードセットを設定します。ポート0x60に書き込まれる次のバイトは、以下のフォーマットのバイトでなければなりません。

ビット0：ポート0x60に設定されている現在のスキャンコードを返す

ビット1：スキャンコードセッ

ト1の設定 ビット2：スキャン

コードセット2の設定 ビット

3：スキャンコードセット3の設

定

その他のビットはすべて0にしてください。

コマンド 0xF3 - オートリピートの遅延とリピート率の設定

このコマンドでは、オートリピートの遅延時間とリピートレートを設定します。ポート0x60に書き込まれる次のバイトは、以下のフォーマットでなければなりません。

ビット0～4：リピートレート 0：約30chars/sec～0x1F：約2chars/sec

ビット5-6 リピートディレイ 00：1/4秒、01：1/2秒、10：3/4秒、11：1秒 他のビットはすべて0にしてください。

リターンコード

ご存知のように、キーボードエンコーダはシステムのオンボードキーボードコントローラと通信します。返される値のはほとんどはスキャンコードですが、時にはエラーを返すこともあります。これらの値は、キーボードデコーダからポート0x60を通じてシステムに送信されます。

戻り値は以下のいずれかです。

Return Codes

Value	Description
0x0	Internal buffer overrun
0x1-0x58, 0x81-0xD8	Keypress scan code
0x83AB	Keyboard ID code returned from F2 command
0xAA	Returned during Basic Assurance Test (BAT) after reset. Also L. shift key make code
0xEE	Returned from the ECHO command
0xF0	Prefix of certain make codes (Does not apply to PS/2)
0xFA	Keyboard acknowledge to keyboard command
0xFC	Basic Assurance Test (BAT) failed (PS/2 only)
0xFD	Diagnostic failure (Except PS/2)
0xFE	Keyboard requests for system to resend last command

オンボード・キーボード・コントローラー・コマンド

これらのコマンドの中には、「A20」の章すでに紹介したものもあります。しかし、ここに挙げたコマンドの多くは新しいもので、中には非常にレベルの低いものもあります。つまり、これらのコマンドの中には、特定のラインをコントロールすることができるものもあります。

がコントローラに接続されています。そのため、コントローラのラインと、キーボードデバイスとのインターフェースを取り上げなければなりませんでした。その他のコマンドでは、コントローラの内蔵RAMを読み書きすることができます。

Command Listing	
Command	Description
Common Commands	
0x20	Read command byte
0x60	Write command byte
0xAA	Self Test
0xAB	Interface Test
0xAD	Disable Keyboard
0xAE	Enable Keyboard
0xC0	Read Input Port
0xD0	Read Output Port
0xD1	Write Output Port
0xE0	Read Test Inputs
0xFE	System Reset
0xA7	Disable Mousr Port
0xA8	Enable Mouse Port
0xA9	Test Mouse Port
0xD4	Write To Mouse
Non Standard Commands	
0x00-0x1F	Read Controller RAM
0x20-0x3F	Read Controller RAM
0x40-0x5F	Write Controller RAM
0x60-0x7F	Write Controller RAM
0x90-0x93	Synaptics Multiplexer Prefix
0x90-0x9F	Write port 13-Port 10
0xA0	Read Copyright
0xA1	Read Firmware Version
0xA2	Change Speed
0xA3	Change Speed
0xA4	Check if password is installed
0xA5	Load Password
0xA6	Check Password
0xAC	Disagnostic Dump
0xAF	Read Keyboard Version
0xB0-0xB5	Reset Controller Line
0xB8-0xBD	Set Controller Line
0xC1	Continuous input port poll, low
0xC2	Continuous input port poll, high
0xC8	Unblock Controller lines P22 and P23
0xC9	Block Controller lines P22 and P23
0xCA	Read Controller Mode
0xCB	Write Controller Mode
0xD2	Write Output Buffer
0xD3	Write Mouse Output Buffer
0xDD	Disable A20 address line
0xDF	Enable A20 address line
0xF0-0xFF	Pulse output bit

これだけの数のコマンドがあるんですね。すべてのコマンドをここで紹介するには、とても長い時間がかかると思いませんか？A20のコマンドについては、すでに「A20」の章で説明しています。このシリーズでは携帯性を重視していますので、上記のような一般的なコマンドのみを取り上げます。しかし、ここで紹介していないコマンドについては、興味のある読者の方に探していただきたいと思います。

次のセクションまでサンプルコードを説明するつもりはありません。むしろ、ここではコマンドそのものを見て、次のセクションからそれらを参照することにします。

コマンド0x20 - コマンドバイトの読み込みとコントローラRAMの読み込み

上の表を見てください。コマンド0x20～0x3Fは、コントローラRAMの読み出しに使われていることに気付きますか？それなのに、コマンド0x20はコマンドバイトの読み出しにも使われています。いったい何が起こっているのでしょうか？

実は、この2つは同じものを指しています。コマンドバイトは、コントローラのRAM内に格納されています。つまり、コマンドバイトを読むときは、コントローラの内蔵RAMから読んでいることになります。いいですか？

コントローラのRAMから読み出す場合、コマンドの最後の6ビットは、RAM内の読み出す場所を示します。一部のMCAシステムでは、RAM内の32のロケーションすべてにアクセスできます。他のシステムでは、0、0x13-0x17、0x1D、0x1Fのバイトにのみアクセスできます。

これらの場所は

オフセット0：コマンドバイト

Offset 0x13 (MCA): nonzero when password is enabled Offset

0x14 (MCA): nonzero when password is matched

オフセット0x16-0x17 (MCA) : パスワード照合時に破棄される2つのマイクロコードを与える オフセット

0x1D。

オフセット0x1F。

ここでは、コマンドバイトの方が重要です。このバイトは、ここに示す特定のビットフォーマットに従います。見た目ほど複雑ではありませんので、ご安心ください。

ビット0：キーボード割り込み

イネーブル0：割り込みを無効

にする

1: キーボード出力のバッファがいっぱいになったときにIRQ 1を送信する

ビット1：マウスインターラップトイネーブル

ISA 未使用

EISA / PS2

0: マウスの割り込みを無効にする

1: マウスの出力バッファがいっぱいになったときにIRQ12を送信する

ビット2：システムフラグ（ステータスレジスタ

のビット2でも可）0: コールドリブート

1: ウォームリブート (BATはすでに完了しています。)

ビット3：キーボードロックを無視する

PS/2: 未使用

- **AT**

0: アクションなし

1: ステータスレジスタのビット4を強制的に1にする（ロックしない）

Bit 4: Keyboard Enable

0: Enable Keyboard

1: クロックラインをLowにしてキーボードを無効にする

ビット5：Mouse Enable

EISAまたはPS/2

0: マウスを有効にする

1: クロックラインをLowにすることでマウスを無効にする

◦ **ISA**

0 : PCモードでは、11ビットコードを使用し、parityチェックとスキャン変換を行う

1 : PCモードでは、8086コードを使用し、parityチェックを行わず、スキャン変換も行わない

ビット6：トランスレーション

0: 翻訳なし

1: キーのスキャンコードを翻訳する。MCAタイプ2のコントローラはこのビットを設定できません

ビット7：未使用、0にすべき

このコマンドが必要になることはないと思いますので、ルーチンは書いていません。

コマンド0x60 - コマンドバイトの書き込みとコントローラRAMの書き込み

コマンドバイト0x60～0x7Fは、上記と非常によく似ており、上記と同じRAMの位置に書き込むことができます。より重要なのは、コントローラのRAMのバイト0（コマンドバイト）を読み出すことで、これはコマンドバイト0x60を送信することで可能です。

コマンド 0xAA - セルフテスト

このコマンドは、コントローラにセルフテストを実行させます。テストの結果は、ポート0x60から読める出力バッファーに返されます。テストが成功した場合は0x55、失敗した場合は0xFCが返されます。

以下は、ルーチンの例です。最初にKYBRD_CTRL_CMD_SELF_TESTコマンド（コマンド0xAA）をキーボードコントローラに送信しているところに注目してください。その後、キーボードコントローラの出力バッファがデータで満たされるのを待ちます。これで、テストが完了したかどうかがわかります。テストが完了すると、出力バッファの結果が0x55であればtrue（テスト成功）、そうでなければfalse（テスト失敗）を返します。

```
//! run self test
bool kkybrd_self_test () {
    //! send command
    kybrd_ctrl_send_cmd (KYBRD_CTRL_CMD_SELF_TEST);

    //! wait for output buffer to be full
    while (1)
        if (kybrd_ctrl_read_status () & KYBRD_CTRL_STATS_MASK_OUT_BUF)
            break;

    //! if output buffer == 0x55, test passed
    return (kybrd_enc_read_buf () == 0x55) ? true : false;
}
```

コマンド 0xAB - インターフェーステスト

このコマンドを実行すると、コントローラとキーボード間のシリアルインターフェースをテストします。テストの結果は、ポート0x60で読める出力バッファーに格納されます。

結果は以下のいずれかになります。0:

成功、エラーなし

1: キーボードのクロックラインが低い

2: キーボードのクロックラインが高止

まり 3: キーボードのデータラインが高

止まり 0xFF: 一般エラー

ご覧のとおり、これらはすべてハードウェアエラーです。エラーが発生した場合は、キーボードを無効にしてリセットすることをお勧めします。それでもダメな場合は、キーボードが故障している可能性があります。

コマンド 0xAD - キーボードの無効化

このコマンドにより、コントローラはキーボードクロックラインをディセーブルにし、コマンドバイトのビット4（キーボードイネーブル）を設定します。コマンドバイトのフォーマットについては、「コマンドバイトの読み出し」の項を参照してください。

つまり、このコマンドはキーボードを無効にします。

キーボードの現在の状態を保存しておくと、システムがキーボードの現在の状態を把握できるようになります。これは、demosのキーボードドライバでは、_kkybrd_disableを通じて行われます。

```
//! disables the keyboard
void kkybrd_disable () {
    kybrd_ctrl_send_cmd (KYBRD_CTRL_CMD_DISABLE);
    _kkybrd_disable = true;
}
```

コマンド 0xAE - Enable Keyboard

このコマンドを実行すると、コントローラはキーボードクロックラインを有効にし、コマンドバイトのビット4（キーボードイネーブル）をクリアします。コマンドバイトのフォーマットについては、「コマンドバイトの読み出し」の項を参照してください。

つまり、このコマンドはキーボードを有効にします。

ここでは、デモから抜粋したルーチンの例を紹介します。いかに簡単かお分かりいただけると思います。)

```

//! enables the keyboard
void kkybrd_enable () {
    kybrd_ctrl_send_cmd (KYBRD_CTRL_CMD_ENABLE);
    _kkybrd_disable = false;
}

```

コマンド 0xC0 - 入力ポートの読み込み

このコマンドは、入力ポート（コントローラのP10～P17ライン）を読み取り、そのバイナリ値をポート0x64から読み取れる出力バッファにコピーします。このポートが持つラインをまだ見ていないので、今から見てみましょう。

Line P10 / Bit 0: キーボードデータ入力、ISAでは未使用

Line P11 / Bit 1: マウスデータ入力、ISAでは未使用 Line

P12 / Bit 2: ISA、EISA、PS/2では未使用

Line P13 / Bit 3: ISA, EISA, PS/2では未使用です。

Line P14 / Bit 4: 0: 512KBマザーボードRAM, 1: 256K RAM

Line P15 / Bit 5 : 0 : 製造用ジャンパー装着、1 : 未装着

Line P16 / Bit 6: 0: CGAディスプレイ 1: MDAディスプレイ

Line P17 / Bit 7: 0: キーボードロック 1: ロックなし

ジャンパーがアクティブな場合、BIOSは無限大の診断ループを実行することがあります。ラインP13とP14は、クロック切り替え用に設定することができます。

これらが複雑に見えても気にしないでください -- 上記を見ると、最近のコンピュータではこのコマンドがあまり役に立たないことがわかるでしょう。ビット0、1、2、3はもう使われません。最近のコンピュータは512KB以上のRAMを搭載しているので、ビット4はほとんど役に立ちません。ビット5は、キーボードテスト用のジャンパーが取り付けられているかどうかをテストするために使用できます（ほとんどのユーザーは行いませんが）。ビット6はビデオアダプターから情報を得られるので必要ありません。ビット7はほとんどのユーザーがキーコードのロックを望まないので、ほとんど必要ありません。なるほど、とても便利なコマンドですね。それは可能ですが、ほとんどのコンピュータには必要ありません。

このコマンドは非常に便利なので、私はこのコマンドのためのルーチンを書かないことにしました。

コマンド 0xD0 - 出力ポートの読み込み

このコマンドは、コントローラの出力ポート（P2）から読み込んで、その結果をポート0x64の出力バッファに格納することを指示します。このコマンドを発行した後にポート0x64から読み込むことで、コントローラの出力ポートのビットを確認することができます。

コントローラの出力ポートは、ちょうどコントローラのP20～P27ラインです（前から覚えていましたか？）このコマンドを実行すると、これらのラインのバイナリ値が出力バッファに格納されます。

出力ポートのピンとその内容については、まだ説明していません。（A20の章で説明しましたが、詳しくは説明していません）なので、ここで説明します。

Line P20 / Bit 0: 0: CPUをリセット、1: 通常動作 Line P21 /

Bit 1: 0: A20ラインを強制、1: 有効 Line P22 / Bit 2: マウスデータ。ISAでは未使用

Line P23 / Bit 3: マウスロック。ISAでは未使用

Line P24 / Bit 4: 0: IRQ 1 not active, 1: IRQ 1 active

Line P25 / Bit 5: 0: IRQ 12 not active, 1: IRQ 12 active

Line P26 / Bit 6: キーボードクロック

Line P27 / Bit 7: キーボードへのデータ

そうです。ビット2と3は、ISA（Industry Standard Architecture）コンピュータ（最近のほとんどのコンピュータ）では使わ

れなくなりました。第4ビットと第5ビット（ラインP24とP25）は、PICラインIR1とIR12でプログラム可能な割り込みコントローラ（PIC）に接続されています。6ビットと7ビットには、現在のキーボードロックとデータ信号が含まれています（ラインがアクティブかどうかに関わらず）。

今のところ、ここに書かれている内容はほとんど役に立たないものばかりですね。これらのビットの多くは、コントローラの現在の動作に関するエレクトロニクスレベルのもので、我々のニーズには役に立ちません。つまり、最初の2行（ビット0と1）を除いては、システムをリセットするか、20番目のアドレスラインを有効/無効にするかを制御します。しかし、ビット0を読んでも意味がありません。このラインはアクティブ（1）でなければならず、正常に動作していることを意味します。これがないとシステムが再起動してしまいます。したがって、ここで役に立つビットはA20ラインだけです。このことは、少なくとも読み出し動作については同じです。

このコマンドがポート0x64で発行された場合、結果のバイトは出力バッファに置かれ、ポート0x60からバイトを読み出すことで読み出すことができます。

すぐにA20を無効にしなければならないという心配はありません。また、キーボードからシステムをリセットする別の方法もあります。このように、このコマンドは私たちのニーズにはほとんど役に立たないので、ルーチンを書かないことにしました。

コマンド 0xD1 - Write Output Port

このコマンドは、出力バッファ（ポート0x60）からバイトをコピーし、コントローラの出力ポートラインにバイトを配置します。これらのラインの説明については、前のセクション（Read Output Port Command）を参照してください。

ほとんどの場合、問題が発生しないように、変更したい特定のビットをビットごとにORし、その他のビットは変更しないようにします。

このコマンドはいくつかの点で便利です。コントローラーが使用するIRQの有効/無効、A20ゲートの有効/無効、さらにビット0を設定することでシステムをリセットすることができます。繰り返しになりますが、変更可能なビットのリストは前のセクションをご覧ください。

コマンド 0xE0 - Read Test Input

このコマンドは、コントローラ上のテストポートラインからバイナリ値を取得し、出力バッファに配置することで、ポート0x60から読み取ることができます。

テストポートは、マイクロコントローラのTEST 0とTEST 1のラインです（本章のコントローラのピンアウト図を参照してください）。この章では、テストポートについて説明していませんので、説明します。

Line TEST 0 / Bit 0 : キーボードクロック（入力）

Line TEST 1 / Bit 1: AT - キーボードデータ(入力) PS/2 - マウスクロック(入力) その他のビットは未定義とみなし、読み出さないでください。

このコマンドはあまり役に立たないかもしれません、コントローラは他の分野でも使用されている可能性があり、テストポートの方が役に立つかもしれないことを覚えておいてください。結局のところ、それはテストのためにあるのです。

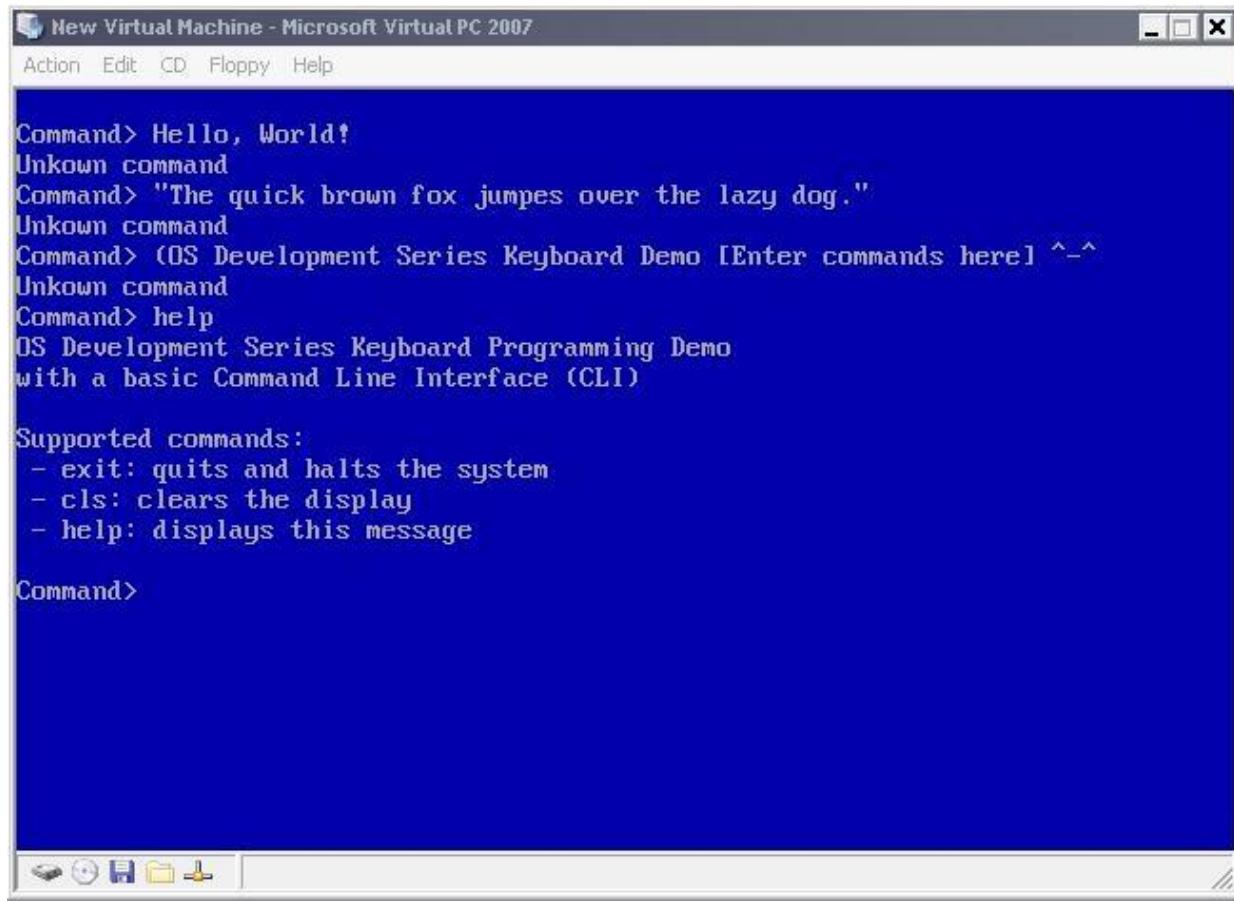
コマンド 0xFE - システムリセット

コントローラの出力ポート（ピンP0）のビット0をパルス化し、CPUをリセットします。これは基本的に、Write Output Portコマンドを送信してビット0をリセットするのと同じことです。システムをきれいにリセットしたい場合は、このコマンドを送信してください。

```
//! reset the system
void kkybrd_reset_system () {
    //! writes 11111110 to the output port (sets reset system line low)
    kybrd_ctrl_send_cmd (KYBRD_CTRL_CMD_WRITE_OUT_PORT);
    kybrd_enc_send_cmd (0xfe);
}
```

すべてのシステムで動作するとは限りませんのでご注意ください。動作するかどうかを確認する簡単な方法は、上記のルーチンの後にプログラムがまだ実行されているかどうかを確認することです：)

Demo



最初のインタラクティブ・デモ

デモダウンロード

これまでのデモの中で最も複雑なものです。前章のコードに加え、キーボードドライバと基本的なコマンドラインインターフェース (CLI) を追加して、より興味深いものにしています。このため、初めてのインタラクティブなデモであり、独自のコマンドで拡張することも可能です。

また、このデモでは、各キーの読み取り間の遅延に使用される `sleep()` ルーチンを追加し、デバッグ出力のコンソールルーチンに画面のスクロールを可能にしています。かっこいいでしょう？

CLI自体は非常にシンプルなものなので、そのコードを紹介するつもりはありません。むしろ、このデモのより重要な部分に焦点を当てたいと思います。

キーボードつなぎ合わせ

このデモのキーボードドライバのルーチンのいくつかをすでに見てきました。キーボードのエンコーダやコントローラとの通信や、有効化、無効化、テスト、LEDの更新、システムのリセットなど、さまざまな重要な機能のルーチンを見てきました。ここまでは良かったのですが、すべてを結びつける重要なディテールがいくつか欠けています。それを見てみましょう。

キーボード。現在の状態を保存する

ご存知のように、キーボードのどのキーもいつでも押せるようになっています。そのため、各キーをスキャンして、キーが押されているかどうかを確認する方法が必要です。ここで良いことに、キーボードエンコーダーはすでにこれを行っているのです。さらに簡単なことに、キーボードエンコーダーはスキャンコードをオンボードのキーボードコントローラーに直接送り、それによってIRQ 1が呼び出されます。

IRQ 1がマスクされていない限り、独自の割り込みハンドラをIRQ 1に設置して、キーボードエンコーダからスキャンコードが送信されるたびに通知を受けることができます。これは何を意味するのでしょうか？私たちの割り込みハンドラは、スキャンコードがキーボードコントローラに送信されるたびに呼び出されます。これはいつでも起こりうることです。

このため、ハンドラ内でスキャンコードをキーボードコントローラにポーリングすることで、スキャンコードを何らかの方法で判定する必要があります。しかし、caps lockキーやnum lockキーのように、キーが押されていない状態では、異なる動作をさせたい場合があります。これらのキーは、押されたときにオンまたはオフになっているはずです。では、shiftのような他のキーはどうでしょうか？これらのキーは、押したままにして、キーを離したときに離す必要があります。

このため、これらのキーの現在の状態と最後に読み取ったスキャンコードを保存し、IRQが戻った後に再び取り出すことができるような方法を考え出す必要があります。これを実現するには

キーボード。割り込み処理

これは重要なことです。キーを押したり離したりするたびに、数バイトのスキャンコードがキーボードコントローラに送られることを覚えていますか？これが発生すると、キーボードコントローラはPIC (Programmable Interrupt Controller) に信号を送り、IRQ 1を発生させます。そうです、読者の皆さん、この信号を受けて、PICはキーボード割り込みハンドラを実行します。

割り込みハンドラの目的は、ドライバの現在の状態を更新することと、スキャンコードをドライバとシステムが使用できる形式に変換して解読することです。そう、それだけのことなんだよ）

割り込みハンドラは、すべてを結びつけるものです。ちょっと大きいので、この文章には載せませんが、皆さんにもぜひ見ていただいて、その動きを確認していただきたいと思います。

キーボード 初期化

キーボードコントローラは、プログラマブル割込みのIRQ1ラインに間接的に接続されていることを覚えていますか？PICを使ってIRQを割込みベクター32 (IRQ 0) からマッピングしたので、IRQ 1は割込みベクター

33. このため、割り込みベクター33を使用するためには、setvectルーチンを使用して割り込みハンドラをインストールする必要があります。

それ以外は非常にシンプルです。kkybrd_set_ledsルーチンを使用して、現在のドライバの状態（グローバルとして保存）をクリアし、LEDをクリアするだけです。

```
///! prepares driver for use
void kkybrd_install (int irq) {

    ///! Install our interrupt handler (irq 1 uses interrupt 33)
    setvect (irq, i86_kybrd_irq);

    ///! assume Basic Assurance test (BAT) test is good
    _kkybrd_bat_res = true;
    _scancode = 0;

    ///! set lock keys and led lights
    _numlock = _scrolllock = _capslock = false;
    kkybrd_set_leds (false, false, false);

    ///! shift, ctrl, and alt keys
    _shift = _alt = _ctrl = false;
}
```

結論

この章はこれで終わりです。システムが面白くなってきたと思いませんか？このデモを発展させて、ある程度使えるようになるかもしれませんね。しかし、現状ではできることが限られています。他のプログラムを走らせて作業を代行することができたら便利だと思いませんか？あるいは、他のファイルをディスクから読み込むこともできたら便利だと思いませんか？ファイルシステム全体の構造を抽象化することは非常に複雑なテーマですが、ここでは1つのディスクからファイルを読み込むことに焦点を当ててみましょう。

しかし、ここで問題が発生する。最低限、ディスクからファイルをロードするためには、まずフロッピーディスクコントローラ (FDC) をプログラムしなければならないのだ。これが次の章のテーマです。お楽しみに：)

次の機会まで。

~マイク

は？お気軽にご連絡ください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。



Chapter 18

Home

Chapter 20

第

18章 ホーム 第20章





オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - FDCプログラミング

by Mike, 2009

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。



8272A Floppy Disk Controller

はじめに

やったーいよいよフロッピーディスクを扱う時がやってきました。この章では、フロッピードライブとフロッピーディスクのプログラミングについて、ほぼすべての知識を網羅しています。

この章のメニューをご紹介します。FDCと

FDDの歴史

ディスクレイア

ウト CHS, LBA

FDD構造 FDC

ハードウェア

FDCとの連携

FDCレジスターとコマンド

歴史

フロッピーディスクコントローラー (FDC) は、フロッピーディスクドライブ (FDD) とのインターフェイスとなるコントローラーです。PCでは通常、このような形で

nec ? pd765 fdc. PS/2はインテル82077Aを使用し、ATはインテル82072Aのマイコンを使用しています。フロッピーディスクドライブ(FDD)は、フロッピーディスクにデータを読み書きすることができる装置です。

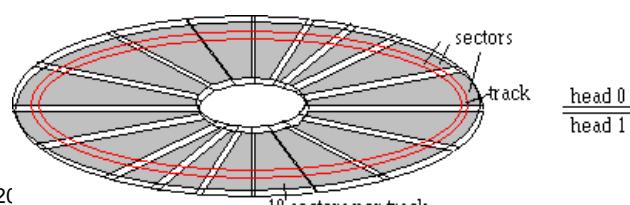
1971年、IBM Direct Access Storage Product ManagerであるAlan Shugartに雇われたDavid L. Nobleは、System/370メインフレーム用の新しいストレージデバイスマルチマットの開発を試みた。IBMは、ICPL (Initial Control Program Load) 用のマイクロコードをリロードする際に、テープドライブよりも小型で高速なものを作りたいと考えていた。ノーブルズのチームは、「ミノー」というコードネームで「メモリーディスク」という製品を開発した。これは、80キロバイトの容量を持つ、読み取り専用の8インチのディスクケットである。1971年に発売され、すべての「System/370」メインフレームに同梱された。

アラン・シュガートがIBMを辞めてメモレックス社に移ると、彼のチームは1972年に初の読み書き両用フロッピーディスクドライブ「Memorex 650」を出荷した。フロッピーディスクは、IBMが8インチ、5と1/4インチ、3と1/2インチのフォーマットを発明した。

ディスク構造

物理的レイアウト

ディスクの構造を理解することは重要です。ここでは、フロッピーディスクのレイアウトを紹介します。



これは、一般的な3-1/2インチフロッピーディスクの物理的なレイアウトです。ここではヘッド1（表側）を見ていますが、セクタは512バイトを表しています。トラックとはセクタの集合体である。

注：1セクタは512バイトで、フロッピーディスクの1トラックは18セクタであることをお忘れなくください。

上の写真を見て、思い出してください。

各トラックは通常、512バイトのセクタに分割されています。フロッピーでは、1トラックに18セクタあります。

シリンドーとは、同じ半径を持つトラックの集まりのことです（上の写真の赤いトラックが1つのシリンドーです）。フロッピーディスクには2つのヘッドがあります（写真に表示されています）。

セクター数は2880。

より深く理解するために、CHSを見てみましょう。次はそれを見てみましょう

シリンダー／ヘッド／セクター (CHS)

セクター

「セクター」とは、簡単に言えば512バイトのグループのこと。つまり、セクター1は、ディスクの最初の512バイトを表しています。

Head

「ヘッド」（またはフェース）は、ディスクの側面を表しています。ヘッド0が表側、ヘッド1が裏側になります。ほとんどのディスクは1面しかないので、ヘッドも1つしかありません（「ヘッド1」）。

トラック

「トラック」とは、ディスクを1周すること。フロッピーディスクの場合、1つのトラックには18セクタが存在する。

シリンド番号は、1枚のディスクのトラック番号を表す。フロッピーディスクの場合は、読み取りを行うトラックを表します。

1トラックに18セクターあります。片面80トラック。

CHSについて

フロッピーディスクのアドレスは、CHSフォーマットを使用している。ディスクの任意の位置から読み書きするためには、FDCにRead/Write Headからディスク上の正確なトラック、シリンダー、セクタに読み書きを行う。

リニア・ブロック・アドレッシング (LBA)

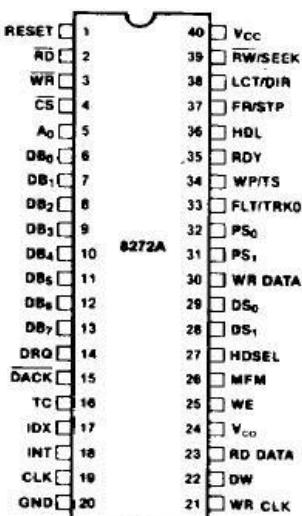
また、LBA (Linear Block Addressing) を用いて、より抽象的なディスクの読み書き方法を提供することもできます。LBAでは、セクタ0～2880までのディスク上の任意のセクタに対して読み書きが可能となる。

フロッピーのインターフェイス

ソフトウェアは、フロッピーディスクコントローラを介してフロッピーディスクドライブを制御することで、フロッピーディスクドライブとのインターフェースを実現しています。ここでは、フロッピーディスクコントローラの種類の違いから、初代の「8272A」を取り上げます。冒頭の画像は、8272Aの代表的なIC（集積回路）コントローラです。ここではこのICを見ていきます。

詳細 82072A フロッピーマイクロコントローラ

8272Aは40ピンのICです。ここではそれを見てみましょう。ここでは、40ピンすべてを簡単に見ていきますが、ここでは、電子工学の分野にまで踏み込んでいくので、詳細には見ません。



これらのピンのほとんどは、コントローラのプログラミングにはあまり役に立ちません。しかし、他のピンは理解するのに重要です。それでは見ていきましょう。わかりやすくするために、すべてのピンを簡単に見ていきます。FDCは間接的に、PIC (Programmable Interrupt Controller) やシステムバス、DMAC (Direct Memory Access Controller) などと通信していることがわかります。

RESET端子 - FDCをアイドル状態にします。すべての出力ラインをLOWにします。Vccピンは、+5Vの電源入力です。

GND端子 - は、グラウンド端子です。

CLK端子 - 典型的な単相8MHz方形波クロック信号。

RD端子 - FDCに現在の動作がリード動作であることを伝える。

WR端子 - 同様の機能を持ちますが、書き込み操作を行います。

これらは、ソフトウェアによるI/Oリード/ライト操作でコントロールバスによって設定されます。

CS端子-チップセレクト

DB0 - DB7ピン - 双方向の8ビットデータバス。システムのプライマリデータバスに間接的に接続されます。

A0端子 - データ/ステータスレジスタセレクト端子。High(1)の場合は、FDCのデータレジスタの内容をデータバスに配置するように指示します。ロー(0)の場合は、ステータス・レジスタの内容をデータ・バスにコピーします。これは、出力データバスピンDB0～DB7を介して行われ、さらに、ソフトウェアで読み取ることができるシステムデータバスを介して行われます。

DRQピン - データダイレクトメモリアクセス(DMA)リクエストピン。このラインがHigh(1)の場合、FDCはDMAリクエストを行っています。

DACKピン - DMAアクノレッジピン。コントローラがDMA転送を行っているときは、このラインはLOW(0)になります。

TC端子 - DMA転送が完了すると、FDCはTerminal Count端子であるTCをH (1) にします。

IDX端子 - FDCがディスクトラックの先頭にあるときにHighになります。

INT端子 - FDCが割り込み要求(IRQ)を送信すると、High(1)になります。のIRQ6に間接的に接続されています。

PIC (Programmable Interrupt Controller)。

RW/Seek端子 - 読み書き両用モードのシークモードを設定する。1: シークモード, 0: リード/ライトモード

LCT/DIR端子 - 低消費電流/方向指定端子。

FR/STP端子 - フォールトリセット/ステップ端子。

HDL端子 - Head Load端子。FDDのRead/Writeヘッドをディスクケットに接触させるコマンドです。

RDY端子 - レディ端子。FDDがデータを送信または受信する準備ができていることを示します。

WP/TS端子 - ライトプロテクト/ツーサイド端子。リード/ライトモードでは、メディアがライトプロテクトされている場合にHighに設定します。シークモードの場合、メディアが両面の場合にHighに設定します。

FLT/TRK0ピン - フォールト/トラック0ピン。Read/Writeモードでは、FDDのフォルトが検出されるとHighになります。

PS0～PS2端子 - 前置補正(Pre-shift)端子です。MFMモード時には、事前補正の状態を書き込みます。

WR DATA端子 - ライトデータ端子

RD DATA端子 - リードデータ端子

DS0 - DS1端子 - ドライブセレクト端子

HDSEL端子 - ヘッドセレクト端子。High(1)の時はFDCがヘッド1にアクセスするように設定します。Lowの時はヘッド0となります。

MFM端子 - Highの場合、FDCはMFMモードで動作します。Low(0)の場合は、FMモードで動作します。

WE端子 - ライトイネーブル端子。

VCO端子 - VCO Sync端子。0であれば、PLL内のVCOを禁止する。1でVCOを有効にします。

DW端子 - データウィンドウ端子。PLLにより生成され、FDDからのサンプルデータに使用さ

れる。WR CLK端子 - ライトクロック

FDCは、DMA (Direct Memory Access) コントローラの有無にかかわらず動作可能です。非DMAモードで動作している場合は、プロセッサとFDCの間でデータバイ特が転送されるたびにIRQ 6が生成されます。DMAモードでは、プロセッサがFDCにコマンドを読み込み、FDCとDMAコントローラの制御下ですべてのデータ転送が行われます。

これは重要なことです。FDCのピンをすべて把握する必要はありません。むしろ、FDCは3つの主要なコントローラと通信していることを覚えておいてください。1つ目は、4つのフロッピーディスクドライブ(FDD)内部コントローラ、プログラマブルインターラップコントローラ(PIC)、ダイレクトメモリアクセス(DMA)コントローラのいずれかです。ソフトウェアは、プロセッサ標準のIN/OUTポートi/o命令によってFDCと通信します。

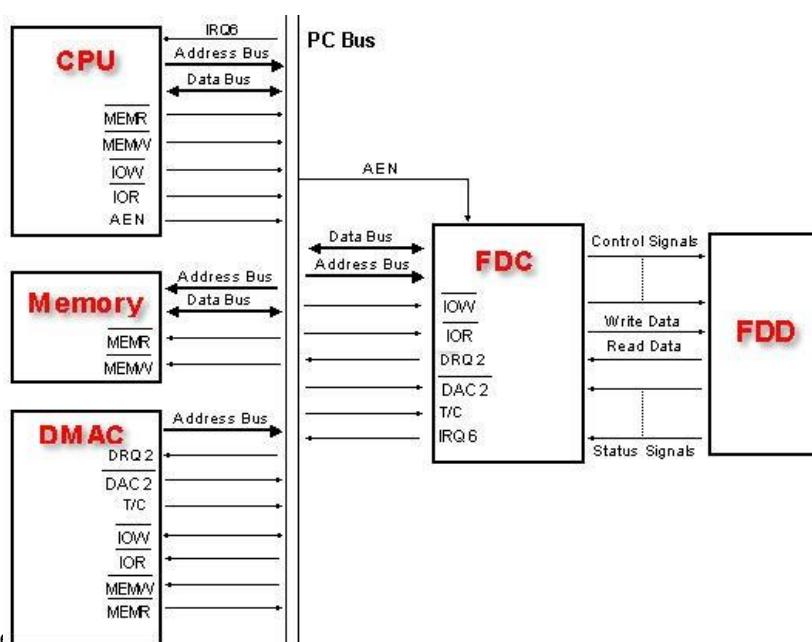
FDCのいくつかのレジスタは、プロセッサのI/Oアドレス空間にマッピングされています。標準的なI/Oポートの読み出しと同様に、入出力動作の間、プロセッサはコントロールバスにREADまたはWRITEラインを設定し、アドレスバスにポートアドレスを設定します。これは、システムバスまたはISA(Industry Standard Architecture)バスで行われます。

最近のハードウェアでは、FDCはISAバスに直接接続されておらず、スーパーI/O ICとして統合されており、スーパーI/OのLow Pin Countバスを通じてプロセッサと通信している。

なるほど！」と思いました。ソフトウェアがFDCと通信する方法はわかりました。PICとDMAの出番は？

上のピンリストを見ると、FDCにはINTというピンがあることがわかります。このラインは、プログラマブルインターラップコントローラのIR 6ラインに間接的に接続されています。FDCは、データのバイトが読み書きできる状態になると、このラインをハイ(1)に引きます。これにより、PICのIR 6ラインもハイに引き出されます。ここからはPICが制御します。他のラインをマスクアウトし、サービスを受けられるかどうかを判断します。プロセッサのインターラップアクノレッジ(INTA)ピンをアクティブにして、プロセッサに割り込みを通知します。プロセッサは、割り込みを処理しても問題ないことを確認すると、INTAラインをリセットして、PICに処理を行うことを許可します。PICは、このIRQが使用するためにマッピングされた(PICの初期化時に設定された)割り込みベクターを配置します。プロセッサはIRQを受け取り、そのアドレスをidtrから取得して、ほら、私たちの割り込みが呼ばれます。

FDCは、DMAモードで動作するようにプログラムすることもできます。DMAは、まだ見たことのないコントローラです。そのため、ここではあまり触れたくありません。しかし、完璧を期すために次の章で説明するかもしれません。FDCはDMAチャンネル2に接続されています。

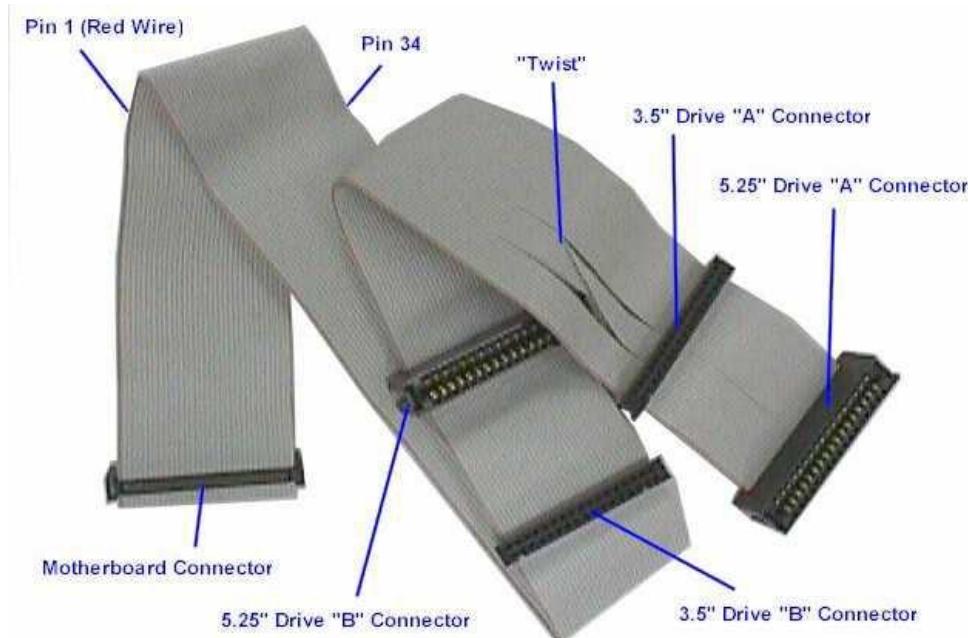


それがFDCのハードウェアのすべてです。FDCは1つのコンピュータシステム内に複数設置することができる。各FDCには最大4台のフロッピーディスクドライブ（FDD）を接続できます。これは重要なことです。FDCと通信する際には、どのFDDに対するリクエストかを選択しなければならないことが多いのです。

フロッピーアンタフェースケーブル

FDCとFDDは、フロッピーアンタフェースケーブルを介して通信します。フロッピーアンタフェースケーブルは、PATA（Parallel ATA）ケーブルの一種です。

ウエスタンデジタル社から発展したIDE（Integrated Drive Electronics）ケーブル。



上記のケーブルにねじれがあることに気づくはずです。それを少し短く説明します。このケーブルには40本のピンがあります。この40本のピンを介して、FDCはケーブルに接続されている異なるFDDと会話することができます。

FDCとの通信に使われるレジスタの中には、コントローラやケーブルの入力ピンを検出できるものがあります。このため、せめてケーブルの40本のラインを少しでも見ておくべきだろう。

Floppy Interface Cable Pins			
Pin	Description	Pin	Description
0	Reset	20	DDRQ
1	Ground	21	Ground
2	Data pin 7	22	I/O Write
3	Data pin 8	23	Ground
4	Data pin 6	24	I/O Read
5	Data pin 9	25	Ground
6	Data pin 5	26	IOCHRDY
7	Data pin 10	27	Cable Select (CS)
8	Data pin 4	28	DDACK
9	Data pin 11	29	Ground
10	Data pin 3	30	Interrupt
11	Data pin 12	31	(No connection)
12	Data pin 2	32	Address 1
13	Data pin 13	33	GPIO_DMA66_Detect
14	Data pin 1	34	Address 0
15	Data pin 14	35	Address 2
16	Data pin 0	36	Chip Select 1
17	Data pin 15	37	Chip Select 3
18	Ground	38	Activity
19	Key or Vcc_in	39	Ground

- 後日、さらに追加されます。

FDCプログラミング

FDCの動作モード

最近のFDCのはほとんどは、オリジナルの8272マイクロコントローラよりも進化しています。後方互換性を確保するために、新しいFDCはコントローラにピンを追加し、特定のモードで動作するときに異なるレジスタと通信できるようにしています。例えば、ステータスレジスタAモードは、コントローラがPC-ATモードで動作しているときにのみアクセスできます。コントローラをリセットすると、コントローラはデフォルトの82077Aモードで動作します。

IRQの待ち受け

FDCはIRQ6を使用していることを覚えていますか？FDCは、リードコマンドやライトコマンドが完了した後、またはモードによっては1バイト転送するごとにバイトを送信します。また、初期化時にコントローラがリセットされるとIRQを送信します。

ここでは、FDCをDMAモードで動作させることにします。基本的には、リード、ライト、シーク、キャリブレートの各コマンドが完了したときと、初期化中のみ割り込みが発生するということです。

しかし、どのような場合でも、コマンドが完了したことを知るために、IRQが発生するのを待つ必要があります。これを実現するには、IRQが発火したときにグローバルを設定し、IRQを待ち、発火したときにグローバルをリセットするirq_waitのような関数を用意します。それでは、早速やってみましょう。まず、IRQです。

```
const int FLOPPY_IRQ = 6;

//! set when IRQ fires
static volatile uint8_t _FloppyDiskIRQ = 0;

void __cdecl i86_flpky_irq () {

    __asm add esp, 12
    __asm pushad
    __asm cli

    //! irq fired
    _FloppyDiskIRQ = 1;

    //! tell hal we are done
    interruptdone( FLOPPY_IRQ );

    __asm sti
    __asm popad
    __asm iretd
}
```

これは、PITのIRQのように簡単そうですね。） そうですね、あとは待つだけですね。

```
//! wait for irq to fire
inline void flpydsk_wait_irq () {

    //! wait for irq to fire
    while ( _FloppyDiskIRQ == 0 )
        ;
    _FloppyDiskIRQ = 0;
}
```

シンプルでいいですね。つまり、読み取りや書き込みなどのコマンドを送るとしたら、単にflpydsk_wait_irq()を呼び出すだけです。これが完了すると、コマンドが終了し、続行しても安全であることがわかります。かっこいいでしょ？）

DMA?

えっ？FDCをDMAモードでプログラミングするの？でも、まだDMAの話をしていないじゃないですか。そうそう、これが問題なんですよ。

私は当初、FDCをNon-DMAモードでプログラムするつもりでした。しかし、多くのエミュレータや一部のハードウェアではDMAモードがサポートされていませんでした。そこで、移植性を確保するためにも、DMA（Direct Memory Access Controller[DMAC]）を利用するのがベストだと判断しました。

しかし、DMAについてはまだ詳しく説明していないため、問題が発生しています。そこで、DMAのインターフェースを説明なしに丸投げするのではなく、3つの基本的なDMAルーチンをハックして、後でより詳細に書き直すことにしました。）

`flpydsk_initialize_dma`は、基本的にDMAが使用するバッファを物理アドレス0x1000～0x10000(64k)に作成します。ディスクからセクターを読み出すとき、DMAはセクターデータをこの場所に置くので、上書きされてしまうので、何もないことを確認してください。他の場所を選んでも構いませんが、いくつかのルールがあります。

バッファは64kの境界を越えることはできません。最高のパフォーマンスを得るためにには、64kの境界に留める必要があります。

書き込み先のメモリ領域は、アイデンティティマップされているか、フレームアドレスがページにマッピングされている必要があります。DMAは常に物理メモリで動作する

デモではバッファに0x1000 + 64kを使用しているので、変更したくない場合はそのままにしておきます。

`dma_read`と`dma_write`は、FDCから送られてきたデータの読み書きを開始するようDMAに指示するだけです。これは、FDCに読み書きを指示したセクターになります。例えば、FDCにセクタの読み出しを指示すると、FDCはDMAにセクタデータを渡し、DMAに設定したバッファ（0x1000）に置かれます。かっこいいでしょ？

```
///! initialize DMA to use phys addr 1k-64k
void flpydsk_initialize_dma () {

    outportb (0x0a, 0x06);    //mask dma channel 2
    outportb (0xd8, 0xff);   //reset master flip-flop
    outportb (0x04, 0);      //address=0x1000
    outportb (0x04, 0x10);
    outportb (0xd8, 0xff);   //reset master flip-flop
    outportb (0x05, 0xff);   //count to 0x23ff (number of bytes in a 3.5" floppy disk track)
    outportb (0x05, 0x23);
    outportb (0x80, 0);      //external page register = 0
    outportb (0xa, 0x02);    //unmask dma channel 2
}

///! prepare the DMA for read transfer
void flpydsk_dma_read () {

    outportb (0xa, 0x06); //mask dma channel 2
    outportb (0xb, 0x56); //single transfer, address increment, autoinit, read, channel 2
```

```

        outportb (0x0a, 0x02); //unmask dma channel 2
    }

//! prepare the DMA for write transfer
void flpydsk_dma_write () {

    outportb (0x0a, 0x06); //mask dma channel 2
    outportb (0x0b, 0x5a); //single transfer, address increment, autoinit, write, channel 2
    outportb (0x0a, 0x02); //unmask dma channel 2
}

```

上記のコードが理解できなくても、心配はいりません。DMAに関するすべてのこととは、次のチュートリアルでDMAを詳しく説明する際に、書き直して説明します。

FDCポートマッピング

FDCには、i86のI/Oアドレス空間にマッピングされた4つの外部レジスターがあります。これらはソフトウェアが標準的なI/O命令でアクセスすることができ
る。これらのレジスターを太字にした。

一部のシステムでは、FDCに提供される外部レジスターの数が、主要な4つのレジスターよりも多い

場合があります。第2のFDCは、通常、I/Oポート0x370～0x377にマッピングされます。

2つの異なるFDCのための2つのポートセットがあるので、この表には両方のポートセットが含まれます。

Floppy Disk Controller Ports			
Port (FDC 0)	Port (FDC 1)	Read/Write	Description
Primary FDC Registers			
0x3F2	0x372	Write Only	Digital Output Register (DOR)
0x3F4	0x374	Read Only	Main Status Register (MSR)
0x3F5	0x375	Read / Write	Data Register
0x3F7	0x377	Read Only	AT only. Configuration Control Register (CCR)
0x3F7	0x377	Write Only	AT only. Digital Input Register (DIR)
Other FDC Registers			
0x3F0	0x370	Read Only	PS/2 only. Status Register A (SRA)
0x3F1	0x371	Read Only	PS/2 only. Status Register B (SRB)
0x3F4	0x374	Write Only	PS/2 only. Data Rate Select Register (DSR)

次のセクションでは、レジスターを少しづつ詳しく見ていきます。とりあえず、重要なものを取り上げます。この章では、完璧を期すために、他のレジスターについても紹介することにするかもしれません。今のところ、上に示した最初の4つのレジスターにのみ焦点を当てます。

これらのコードはすべて、この章の最後にあるデモに含まれていることを忘れないでください。

```

FLPYDSK_DOR      =      0x3f2,
FLPYDSK_MSR      =      0x3f4,
FLPYDSK_FIFO     =      0x3f5, //data register
FLPYDSK_CTRL     =      0x3f7

```

レジスター

ステータスレジスタA (SRA) (PS2モードのみ)

このレジスタを知っている必要はありません。ここにあるのは完全性のためだけです。

このレジスタは、コントローラのいくつかのインターフェースピンの状態を監視するリードオンリーのレジスタです。コントローラがPC-ATモードの場合
はアクセスできません。これは読み取り専用のレジスタです。

このレジスタの正確なフォーマットは、コントローラのモデルによって異なります。

- ビット0 DIR
- ビット1 WP
- ビット2 INDX
- ビット3 HDSEL
- ビット4 TRKO
- ビット5 STEPフリップ/フロップ
- ビット6 DRV2
- ビット7 INTERRUPTラインの状態（割込み保留）

警告これらのビットは、コントローラのモデルによって変わることがあります。

ステータスレジスタB (SRB) (PS/2モードのみ)

このレジスタを知っている必要はありません。ここにあるのは完全性のためだけです。

上記のレジスタと同様に、FDCのいくつかのラインの状態を監視することができます。FDCがPC-ATモードの時はアクセスできません。このレジスタは読み取り専用です。

ビット0 MOT EN0 (モータ・イネーブル0) ビット1 MOT EN1 (モータ・イネーブル1) ビット2 WEフリップ/フロップ
 Bit 3 リードデータ (RDDATA) フリップ/フロップ Bit 4 ライトデータ (WRDATA)
 フリップ/フロップ Bit 5 ドライブセレクト 0
 ビット6 未定義、常に1
 ビット7 未定義、常に1

警告 これらのビットは、コントローラのモデルによって変わることがあります。

このレジスタが複雑に見えても気にしないでください。本シリーズでは使用しませんのでご了承ください。

データレートセレクトレジスタ (DSR)

このレジスタを知っている必要はありません。ここにあるのは完全性のためだけです。

このレジスタは、ドライブ制御信号のタイミングを変更することができる書き込み専用のレジスタです。I/Oポート0x3f4 (FDC 0) または0x374 (FDC 1) に書き込むことで使用されます。

これは8ビットのレジスタです。以下のようなフォーマットになっています。

ビット0 DRATE
 SEL0 ビット1
 DRATE SEL1 ビット2 PRE-COMP 0
 ビット3 PRE-COMP 1
 ビット4 PRE-COMP 2
 Bit 5 Must be 0
 Bit 6 POWER DOWN : 内部クロックを停止し、内部発振器を停止する
 ビット7 S/W RESET。内部振動子のリセット

PRE-COMP 0～PRE-COMP 2は少し複雑です。これらは、フロッピードライブなどの磁気メディアで発生するビットシフトに対して、WRDATA出力ピンを調整します。前置補正の遅延を調整するには、これらのビットを以下のいずれかに設定します。

000 デフォルト (250～500Kbps : 125ns, 1Mbps : 41.67ns)
 110 250 ns
 101 208.33 ns
 100 166.67 ns
 011 125 ns
 010 83.34 ns
 001 41.67 ns
 111 無効化

DRATE SEL0 - DRATE SEL1でデータレートを設定します。有効な値を以下に示します。

10 00 500 Kbps
 250 Kbps
11 01 300 Kbps

1Mbps

警告 データレートをドライブの許容範囲を超えて設定すると、エラーが発生する可能性があります。

デジタル出力レジスタ (DOR)

Yey! 最初の便利なレジスター! これは知っておいて損はない。

これは書き込み専用のレジスタで、FDDのモーター制御、動作モード (DMA, IRQ), リセット、ドライブなど、FDCのさまざまな機能を制御することができます。フォーマットを持っています。

00 ビット0-1 DR1, DR2
 01 - ドライブ0
 ◦ 02 - Drive 1
 ◦ 10 - Drive 2
 ◦ 11 - Drive 3

- 1 - コントローラのリセット
 - 2 - Controller enabled
- 0 ビット3 モード
- IRQチャンネル1 -
 - DMAモード
 - 1 ビット4~7 モーターコントロール
 - (ドライブ0~3) 0 - ドライブのモーター停止
 - 2 - Start Motor for drive

これは簡単なことです。基本的には、FDCの機能を制御するコマンドを送信する際には、どのドライブ用か（1台のFDCが4台のFDDと通信できることを覚えておいてください！）、コントローラのリセット状態、動作モード（FDCはDMAとIRQの両方のモードで動作できることを覚えておいてください！）、特定のFDD内部モーターの状態を選択するためのビットパターンを構築するだけです。

以下にその例を示します。最初のフロッピーディスクドライブ（FDD 0）のモーターを起動させたいとします。FDDのモーターを起動させるには、FDDに対して読み書きの操作を行う前に行う必要があります。モーターを起動させるには、起動または停止させたいドライブに対応するビット（4-7）を設定します。他のビットを0にしておくと、通常の動作（IRQモード、コントローラのリセット）になります。DORがプロセッサのi/oアドレス空間のポート0x3f2にマッピングされていることを知っているので、これは非常に簡単になります。まず、読みやすさを向上させるために、レジスタのビットマスクを作成します。これらのコードはすべて、このチュートリアルの最後にあるデモにも含まれていることを覚えておいてください。

```
enum FLPYDSK_DOR_MASK {...
```

上記のビットマスクを使用すると、設定したい異なるビットをビットごとにORするだけよいのです。たとえば、フロッピーディスクドライブ0のモーターを起動するには、次のようにします。

```
outportb (FLPYDSK_DOR, FLPYDSK_DOR_MASK_DRIVE0_MOTOR | FLPYDSK_DOR_MASK_RESET);
```

FLPYDSK_DORは先ほど0x3f2と定義しましたが、これはDOR FDCレジスタのi/oアドレスであることを思い出してください。また、上記はコントローラをリセットします。

この同じモーターをオフにするには、モータービットを設定せずに同じコマンドを送信します。

```
outportb (FLPYDSK_DOR, FLPYDSK_DOR_MASK_RESET);
```

警告 モーターの起動には時間がかかります。内蔵のFDDモーターは機械式であることを忘れないでください。他の機械装置と同様に、ソフトウェアの動作速度よりも遅くなる傾向があります。このため、FDDモーターを起動する際には、読み取りまたは書き込みを行う前に、必ず少し時間を置いてから回転させてください。

DORは書き込み専用のレジスタです。これを実現するために、ルーチンを作成します。

```
void flpydsk_write_dor (uint8_t val) {
    //! write the digital output register
    outportb (FLPYDSK_DOR, val);
}
```

次の重要な登録に移りましょう。

メイン・ステータス・レジスター (MSR)

メイン・ステータス・レジスター (MSR) は、特定のビットフォーマットに従っています。これは予想外だったでしょう。さて、話を元に戻しましょう (シャレです)。MSRのフォーマットは次のとおりです。

ビット0 - FDD 0 : FDDがシークモードでビジーの場合は1	ビット1 - FDD 1 : FDDがシークモードでビジーの場合は1
ビット2 - FDD 2 : FDDがシークモードでビジーの場合は1	ビット3 - FDD 3 : FDDがシークモードでビジーの場合は1
ビット4 - FDC Busy; ReadまたはWriteコマンド進行中 1: 選択したFDDがビジー状態 0: Not busy	ビット5 - 非DMAモードのFDC 0: DMAモードのFDC
ビット6 - DIO: FDC ICとCPU間のデータ転送の方向 1: FDCはCPUからのデータを期待している 0: FDCはDMAモードではない	ビット7 - RQM。データレジスタはデータ転送の準備ができている 1: データレジスタレディ 0: データレジスタは準備ができていない

このMSRはシンプルなものです。FDCとディスクドライブの現在のステータス情報を含んでいます。コマンドを送信したり、FDDから読み出したりする前に、常にFDCの現在のステータスをチェックして、準備ができていることを確認する必要があります。

ここでは、このMSRがビジー状態かどうかを読み取る例を示します。まず、コードで使用されるビットマスクを定義します。上の図のようなフォーマットになっていることに注目してください。

```
FLPYDSK_MSR_MASK_DRIVE1_POS_MODE      =      1,      //00000001
FLPYDSK_MSR_MASK_DRIVE2_POS_MODE      =      2,      //00000010
FLPYDSK_MSR_MASK_DRIVE3_POS_MODE      =      4,      //00000100
FLPYDSK_MSR_MASK_DRIVE4_POS_MODE      =      8,      //00001000
FLPYDSK_MSR_MASK_BUSY                 =     16,      //00010000
FLPYDSK_MSR_MASK_DMA                  =     32,      //00100000
FLPYDSK_MSR_MASK_DATAIO               =     64,      //01000000
FLPYDSK_MSR_MASK_DATAREG              =    128,      //10000000
```

簡単でしょうか？では、FDCがビジー状態（BUSYフラグがセットされている）かどうかをテストしてみましょう。FLPYDSR_MSRがMSRのi/oポートアドレスである0x3f4であることを知っているので、必要なことは次のとおりです。

```
if ( inportb (FLPYDSK_MSR) & FLPYDSK_MSR_MASK_BUSY )
    //! FDC is busy
```

読み書きのコマンドを送るときは、このビットが0になるまで待てばいいのです。

読みやすくするために、これをルーチンに隠すことにしたので、ここに紹介します。このルーチンは、FDCのステータスを返すだけです。

```
uint8_t flpydsk_read_status () {
    //! just return main status register
    return inportb (FLPYDSK_MSR);
}
```

テープドライブレジスター (TDR)

このレジスタを知っている必要はありません。ここにあるのは完全性のためだけです。

このレジスタは、特定のドライブの初期化時に、そのドライブにテープ・ドライブ・サポートを割り当てることができます。このレジスタはリード／ライトレジスタで、サイズは8ビットです。しかし、最初の2ビットのみが定義されています。ドライブ0はフロッピーブートデバイス用に予約されているため、選択することはできません。そのため、以下のビットリストには含まれていません。

- 00: None.
- 01: ドライブ1
- 10: ドライブ2

- 11: ドライブ3

このレジスターは、ハードウェアリセットのみでリセットされます。ソフトウェアのリセットでは効果がありません。テープドライブに詳しくなくても心配りません。このレジスタは私たちには適用されませんし、このシリーズでは使用しません。このレジスタは我々には適用されず、このシリーズでは使用されません。)

データレジスタ

これは、8ビットまたは16ビットのリード／ライトレジスタです。レジスタの実際のサイズは、コントローラのタイプによって異なります。すべてのコマンドパラメタとディスクデータの転送は、データレジスタへの読み出しと書き込みを行います。このレジスタは、特定のビットフォーマットに従わず、一般的なデータに使用されます。I/Oポート0x3f5(FDC 0)または0x375(FDC 1)からアクセスできます。

注：このレジスタを読み書きする前に、まずマスター・ステータス・レジスタ (MSR) のステータスを読み、そのレジスタが有効であることを確認する必要があります。

覚えておいてください。すべてのコマンドバイトとコマンドパラメータは、このレジスタを介してFDCに送信されます。以下のコマンドセクションでこの例を見るることができますので、まだあまり気にしないでください。

無効なコマンドが発行された場合、データレジスタから返される値は0x80です。

以下のルーチンは、このレジスタから読み取り、デモで使用されます。このルーチンは、データ・レジスタが読み書き可能な状態になるまで待機し、その後、データの読み取り (read_data関数) または書き込み (send_command関数) を行います。

```
void flpydsk_send_command (uint8_t cmd) {
    //! wait until data register is ready. We send commands to the data register
    for (int i = 0; i < 500; i++)
        if ( flpydsk_read_status () & FLPYDSK_MSR_MASK_DATAREG )
            return outportb (FLPYDSK_FIFO, cmd);
}

uint8_t flpydsk_read_data () {
    //! same as above function but returns data register for reading
    for (int i = 0; i < 500; i++)
        if ( flpydsk_read_status () & FLPYDSK_MSR_MASK_DATAREG )
            return inportb (FLPYDSK_FIFO);
}
```

デジタル入力レジスタ (DIR)

このレジスタを知っている必要はありません。ここにあるのは完全性のためだけです。

さて、デジタル出力レジスタ (DOR) がありましたので、これは予想していたことだと思いますが:) このレジスタは、コントローラのすべての動作モードで読み取り専用です。PC-ATモードではビット7のみが定義され、他のビットは未定義であり、使用してはいけません。他の動作モードでは、ビット7は定義されません。

ビット7（DSK CHG）は、FDCのDSK CHG端子をモニターします。本章冒頭のピン配置を見ると、DSK CHG端子がないことがわかります。これは、FDCの最新モデルと初代モデルの違いに関係しています。新しいモデルでは、DMA GATEやDRATE SEL0/1など、FDCの新しいピンを監視するために、このレジスタの異なるビットが追加・変更されています。このレジスタの値は、FDCの動作モードに固有のものです。

なお、このレジスタのビットは、モデルによって変更されることがあります。

コンフィグレーション・コントロール・レジスタ（CCR）

PC/ATモードでは、このレジスターはデータレートセレクトレジスター（DSR）と呼ばれ、最初の2ビット（ビット0=DRATE SEL0、ビット1=DRATE SEL1）のみが設定されています。もう一度見てみましょう…。

ビット2は、モデル30/CCRモードではNOPRECであり、機能はありません。その他のビットは未定義で、コントローラによって変化する可能性があります。他のレジスタと同様に、このレジスタに書き込めるようにルーチンを作成しました。

```
void flpydsk_write_ccr (uint8_t val) {  
  
    //! write the configuration control  
    outportb (FLPYDSK_CTRL, val);  
}
```

コマンド

アブストラクト

コマンドは、FDCに接続されたFDDを制御して、読み出しや書き込みなどの異なる動作を行うために使用される。コマンドは、データバス (D0-D7) 端子を介して、データレジスタに書き込み操作で書き込まれます（コントロールバスにはIOおよびWRITEコントロールラインが設定されています）。

警告 コマンドやパラメータバイトを送信する前に、まずMSR (Main Status Register) のビット7をテストして、データレジスタがデータを受信する準備ができていることを確認してください。

13個（コントローラによってはそれ以上）のコマンドがあります。各コマンドのサイズは1~9バイトです。FDCは、最初のコマンドバイトから何バイトを期待するかを知っています。つまり、最初のバイトは、FDCに何をしてほしいかを伝える実際のコマンドである。FDCは、このコマンドからさらに何バイトを期待するかを知っています（コマンドパラメタ）。

コマンドは、トラックの片方のヘッドでのみ動作します。両方のヘッドで動作させたい場合は、Multiple Track Bitを設定する必要があります。これらのコマンドの多くは、ビットフォーマットに従っています（後述します）。ここからが複雑なのです。

コマンドバイトは、下位4ビットのみを実際のコマンドに使用します（それ以上の場合もあります）。これらのコマンドバイトの上位ビットは、コマンドのオプション設定用です。私はこれを拡張コマンドビットと呼んでいますが、正式な名称はありません。これらのビットの中には、私たちが使用する必要のあるほとんどのコマンドに共通するものがいくつかあります。これらのビットについては、後ほどコマンドバイトの中で説明します。

さて、まずはコマンドリストを見てみましょう。次に、それぞれのコマンドを個別に見てていきます。どれもコマンドバイトの最初の4ビットしか使っていないことに注目してください。

FDC_CMD_READ_TRACK	=	2,
FDC_CMD_SPECIFY	=	3,
FDC_CMD_CHECK_STAT	=	4,
FDC_CMD_WRITE_SECT	=	5,
FDC_CMD_READ_SECT	=	6,
FDC_CMD_CALIBRATE	=	7,
FDC_CMD_CHECK_INT	=	8,
FDC_CMD_WRITE_DEL_S	=	9,
FDC_CMD_READ_ID_S	=	0xa,
FDC_CMD_READ_DEL_S	=	0xc,
FDC_CMD_FORMAT_TRACK	=	0xd,
FDC_CMD_SEEK	=	0xf

FDCにコマンドを送るには、FIFOと呼ばれるデータレジスタにコマンドを書き込まなければならないことを覚えておいてください。そのためには、まずMSRのビットをチェックしてデータレジスタの準備が整うのを待つ必要があります。flpydsk_read_status()はMSRからの値を返すだけなので、もっと簡単なメソッドの中にすべてを隠すことができます。

```
void flpydsk_send_command (uint8_t cmd) {  
  
    //! wait until data register is ready. We send commands to the data register  
    for (int i = 0; i < 500; i++)  
        if ( flpydsk_read_status () & FLPYDSK_MSR_MASK_DATAREG )  
            return outportb (FLPYDSK_FIFO, cmd);  
}
```

拡張コマンドビット

これらのコマンドの中には、コマンドが実行される前に数バイトを渡さなければならないものがあります。また、複数のバイトを返すものもあります。読みやすくするために、すべてのコマンド、フォーマット、パラメタのバイトを表にしました。各コマンドには、説明とサンプル・ルーチンが付いています。

さて、拡張コマンドビットの話をしたときに、上のコマンドが4ビットしかないことを覚えていますか？この上位4ビットは、さまざまな用途に使用できます。•

例えば、Write SectorコマンドのフォーマットはM F 0 0 0 1 1 0で、最初の4ビット（0 1 1 0）がコマンドバイト、上位4ビット（M F 0 0）が各種設定を表しています。

Mはマルチトラック、Fはコマンドの動作密度モードを選択するための設定です。

ここでは、一般的なビット

のリストを紹介します。

M - マルチトラック・オペ

レーション

- 0 : シリンダーの片側のトラックで動作 1 : シリンダーの両側のトラックで動作
- F - FM/MFMモード設定
 - 0 : FM（単密度）モードで動作 1 : MFM（倍密度）モードで動作
- S - スキップモード設定
 - 0 : 削除済みデータのアドレスマークをスキップ 1 : 削除済みデータのアドレスマークをスキップする
- HD - ヘッドナンバー

DR0 - DR1 - ドライブ番号ビット (2ビットで最大4台のドライブに対応)

M、F、Sの各ビットは多くのコマンドに共通しているので、それらをまとめて表示することにしました。これらのビットを設定するには、これらの設定と使用したいコマンドをビット和するだけです。

```
enum FLPYDSK_CMD_EXT {
    FDC_CMD_EXT_SKIP      = 0x20, //00100000
    FDC_CMD_EXT_DENSITY   = 0x40, //01000000
    FDC_CMD_EXT_MULTITRACK = 0x80 //10000000
};
```

GAP 3

GAP3とは、物理ディスク上のセクタ間のスペースのこと。GPL (Gap Length) の一種です。

```
enum FLPYDSK_GAP3_LENGTH {
    FLPYDSK_GAP3_LENGTH_STD = 42,
    FLPYDSK_GAP3_LENGTH_5_14= 32,
    FLPYDSK_GAP3_LENGTH_3_5= 27
};
```

一部のコマンドでは、GAP3コードを通す必要があるので、そこは注意が必要です。)

セクタあたりのバイト数

一部のコマンドでは、セクタごとのバイト数を入力する必要があります。しかし、これらのバイト数は任意の大きさにすることはできず、常に計算式に従います。

$2^n * 128$, where ^ denotes "to the power of"

n は0~7の数字です。 $2^7 * 128 = 16384$ (16Kバイト) なので、7以上にはならない。FDCでは、1セクタあたり最大16Kバイトまで選択することが可能ですが。しかし、ほとんどのドライブはこれをサポートしていないかもしれません。

このリストには、最も一般的なものが含まれています。

```
enum FLPYDSK_SECTOR_DTL {
    FLPYDSK_SECTOR_DTL_128  = 0,
    FLPYDSK_SECTOR_DTL_256  = 1,
    FLPYDSK_SECTOR_DTL_512  = 2,
    FLPYDSK_SECTOR_DTL_1024 = 4
};
```

…だから、コマンドでセクタあたりのバイト数を要求されたら、512とは言わない！むしろ、FLPYDSK_SECTOR_DTL_512、つまり2とする。

コマンドにパラメタを渡す方法

思い出せば、多くのコマンドではパラメーターを渡す必要があります。パラメーターを渡すには、コマンドが送られてきたのと同じ方法でパラメーターを送るだけです。例えば、specifyコマンドでは、2つのパラメタを渡す必要があります。これがないとコマンドが起動しませんので…。

```
f1pydsk_send_command (FDC_CMD_SPECIFY);
f1pydsk_send_command (data);
f1pydsk_send_command (data2);
```

それが全てです。)

コマンドから戻り値を取得する方法

プログラミングの関数のように戻り値を無視できるのではなく、FDCでは戻り値を何らかの形で処理する必要があります。もちろん、無視してもいいので

ですが、FDCから受け取る必要があります。それが終わるまでFDCはそれ以上のコマンドを許さない。

コマンドがデータを返す場合、そのデータはFIFO（データレジスタ）に--1つずつ--返されます。そのため、読むためには、FIFOを継続的に読み込んで、返されたデータをすべて取得する必要があります。

注：コマンドがデータを返す場合は、待つ必要のある割り込みが送られてきます。これにより、コマンドが終了し、FIFOから戻り値を読んでも問題ないことを知ることができます。

戻り値の良い例として、read sectorsコマンドがあります。このコマンドは、完了を知らせるためにIRQを待つ必要があり、7バイトを返します。そのため、返されたデータバイトをすべて読み出すには、データレジスタから1つずつ読み出す必要があります。

```
for (int j=0; j<7; j++)
    flpydsk_read_data ();
```

もちろん、エラーチェックのためには、実際にいくつかの戻り値を確認する必要があります。

ライトセクター

形式です。M F 0 0 0 1 1 0

Paramaters:

x x x x x HD DR DR0 シリ

ンダー

- Head
- セクター番号 セ
- クターサイズ ト
- ラック長 GAP3
- の長さ データ長

戻る。

リターンバイト 0 :

ST0 リターンバイト

1 : ST1 リターンバイ

ト 2 : ST2

Return byte 3: 現在のシリンドー

Return byte 4: 現在のヘッド Return

byte 5: セクター番号 Return byte 6:

セクターサイズ

このコマンドは、FDDから1セクタを読み出すものです。セクター内の1バイトごとに、FDCは割り込み6を発行し、ディスクから読み込んだバイトをデータレジスタに入れて、読み込めるようにします。

セクターを読む

形式です。M F S 0 0 1 1 0

Paramaters:

x x x x x HD DR1 DR0 = HD=ヘッド DR0/DR1=ディス

クシリンドー

- Head
- セクター番号 セ
- クターサイズ ト
- ラック長 GAP3
- の長さ データ長

戻る。

リターンバイト 0 :

ST0 リターンバイト

1 : ST1 リターンバイ

ト 2 : ST2

Return byte 3: 現在のシリンドー

Return byte 4: 現在のヘッド Return

byte 5: セクター番号 Return byte 6:

セクターサイズ

このコマンドは、FDDから1セクタを読み出すものです。セクター内の1バイトごとに、FDCは割り込み6を発行し、ディスクから読み込んだバイトをデータレジスタに入れて、読み込めるようにします。

以下は、デモで使用したルーチンです。まず、DMAを設定して読み出し動作の準備をします。その後、リードセクターコマンド(FDC_CMD_READ_SECT)を実行し、コマンドM、F、Sの各ビットを設定します(マルチトラックリード、倍密度、削除されたアドレスマークのスキップ。これらの一覧は上記を参照してください)。

その後、すべてのコマンドパラメタを渡して、読み出しコマンドを開始します。セクターサイズはFLPYDSK_SECTOR_DTL_512 (Bytes per sector) で、

値は2です(詳細は上記のBytes per sectorの項を参照してください)。次のパラメーターはGAP3の長さです。ここでは、標準的な3-1/2インチフロッピーデ

ィスクのGAP3長(FLPYDSK_GAP3_LENGTH_3_5、27)を渡します。

データ長パラメータバイトは、セクターサイズが0の場合のみ有効で、それ以外の場合は0xffとします。このコマンドは完了後にIRQを送信するため、IRQを待つ必要があります。

```
void flpydsk_read_sector_imp (uint8_t head, uint8_t track, uint8_t sector)
{
    uint32_t st0, cyl;

    //! set the DMA for read transfer
    flpydsk_dma_read ();

    //! read in a sector
    flpydsk_send_command (
        FDC_CMD_READ_SECT | FDC_CMD_EXT_MULTITRACK |
        FDC_CMD_EXT_SKIP | FDC_CMD_EXT_DENSITY);
    flpydsk_send_command ( head << 2 | _CurrentDrive );
    flpydsk_send_command ( track);
    flpydsk_send_command ( head);
    flpydsk_send_command ( sector);
    flpydsk_send_command ( FLPYDSK_SECTOR_DTL_512 );
    flpydsk_send_command (
        ( ( sector + 1 ) >= FLPY_SECTORS_PER_TRACK )
        ? FLPY_SECTORS_PER_TRACK : sector + 1 );
    flpydsk_send_command ( FLPYDSK_GAP3_LENGTH_3_5 );
```

```

    flpydsk_send_command ( 0xff );

    //! wait for irq
    flpydsk_wait_irq ();

    //! read status info
    for (int j=0; j<7; j++)
        flpydsk_read_data ();

    //! let FDC know we handled interrupt
    flpydsk_check_int (&st0,&cyl);
}

```

...IRQが発火した後、7つのリターンバイトをすべて読み込みます。そして、flpydsk_check_int()でSENSE_INTERRUPTコマンドを送り、FDCに割り込みを処理したこと伝えます。(後述の「インタラプトステータスの確認」の項を参照してください)。

待って....。データはどこにあるの?上のコマンドを見ると、FDCにデータをどこに置くかを指示していません。これは面白い問題だと思いませんか?

FDCの動作モードにもよりますが、Non-DMAモードでは、1バイトごとにIRQ6を発射します。ディスクから読み込んだデータのバイトはFIFOに入っています。DMAモードでは、データをDMAに渡し、DMAはそのデータをバッファ(DMAに置くように指示した場所)に入れます。

つまり、今回のケースでは、DMAバッファを0x1000に設定しましたよね?上記のルーチンを呼び出した後、セクタデータは0x1000になります! かっこいいでしょう? DMAに別のアドレスを与えることで、その位置を変更することができます。

ドライブデータの修正/指定

フォーマット 0 0 0 0 0 0 1 1 パラ

メーター

S S S S H H H H - S=ステップレート H=ヘッドアンロード時間

H H H H NDM - H=Head Load Time NDM=0(DMA Mode)または1(DMA Mode)の場合 Return:

なし

このコマンドは、FDCに接続されているメカニカル・ドライブの制御情報をFDCに伝えるために使用されます。このコマンドの操作を簡単にするために、そのためのルーチンを書いてみましょう。

```

void flpydsk_drive_data (uint32_t stepr, uint32_t loadt, uint32_t unloadt, bool dma )

{
    uint32_t data = 0;

    flpydsk_send_command (FDC_CMD_SPECIFY);

    data = ( (stepr & 0xf) << 4) | (unloadt & 0xf);
    flpydsk_send_command (data);

    data = (loadt) << 1 | (dma==true) ? 1 : 0;
    flpydsk_send_command (data);
}

```

ステータスの確認

フォーマット 0 0 0 0 0 1 0 0

パラメーター

x x x x x x HD

DR1 DR0 戻ります。

バイト 0: ステータス・レジスタ 3 (ST3) このコマンドは、ドライブ・ステータスを返します。

ドライブのキャリブレーション

フォーマット 0 0 0 0 0 1 1 1

パラメーター

x x x x x x 0 DR1 DR0 戻り

値: なし

このコマンドは、リード/ライトヘッドをシリンドラ0に位置付けるために使用します。完了後、FDCは割り込みを発行します。ディスクのトラック数が80以上の場合は、このコマンドを数回発行する必要があります。このコマンドを発行した後は、必ず正しいトラックにあるかどうかを確認してください(Check Interrupt Statusコマンド)。

コマンドを実行した後、まだシリンドラ0に到達していない場合は、再度コマンドを実行します。0番のシリンドラを見つけたら、モーターをオフにして成功を返します。10回やっても見つからなければ中止します。

このコマンドを実行する際には、モーターが動作していることを確認する必要があることに注意してください。また、SENSE_INTERRUPTコマンド(flpypdsk_check_int()コール)を使って、現在のシリンドラの状態を確認していることにも注目してください。

```
int flypydsk_calibrate (uint32_t drive)
{
    uint32_t st0, cyl;
    if (drive >= 4)
        return -2;
    //! turn on the motor
```

```

    flpydsk_control_motor (true);

    for (int i = 0; i < 10; i++) {

        //! send command
        flpydsk_send_command ( FDC_CMD_CALIBRATE );
        flpydsk_send_command ( drive );
        flpydsk_wait_irq ();
        flpydsk_check_int ( &st0, &cyl );

        //! did we fine cylinder 0? if so, we are done
        if (!cyl) {

            flpydsk_control_motor (false);
            return 0;
        }
    }

    flpydsk_control_motor (false);
    return -1;
}

```

割り込みステータスの確認

フォーマット 0000100

0 パラメーター None

Return:

バイト0: ステータスレジスタ0 (ST0)

バイト1: 現在のシリンダー

割り込み復帰時のFDCの状態の情報を確認するためのコマンドです。

```

void flpydsk_check_int (uint32_t* st0, uint32_t* cyl)

{ flpydsk_send_command (FDC_CMD_CHECK_INT);

    *st0 = flpydsk_read_data ();
    *cyl = flpydsk_read_data ();
}

```

シーク／パークヘッド

フォーマット: 00000111 パラ

メーター

xxxxxx HD DR1 DR0 - HD=ヘッド DR1/DR0=ドライ

ブシリンダー

戻る。なし

このコマンドは、リード／ライトヘッドを特定のシリンダーに移動するために使用します。calibrate コマンドと同様に、このコマンドを複数回送信する必要があります。試行錯誤のたびに check_int () を呼び出して現在のシリンダーを取得していることに注目してください。その後、現在のシリンダーが探しているシリンダーであるかどうかをテストします。もしそうでなければ、もう一度試します。もしそうであれば、成功を返します。

```

int flpydsk_seek ( uint32_t cyl, uint32_t head )

{ uint32_t st0, cyl0;

    if (_CurrentDrive >= 4)
        return -1;

    for (int i = 0; i < 10; i++) {

        //! send the command
        flpydsk_send_command (FDC_CMD_SEEK);
        flpydsk_send_command ( (head) << 2 | _CurrentDrive);
        flpydsk_send_command (cyl);

        //! wait for the results phase IRQ
        flpydsk_wait_irq ();
        flpydsk_check_int ( &st0, &cyl0 );

        //! found the cylinder?
        if ( cyl0 == cyl )
            return 0;
    }

    return -1;
}

```

無効なコマンド

無効なコマンドがFDCに送られてきた場合、FDCはそれを無視してスタンディモードになります。

FDCのリセット

コントローラーの無効化

DOR RESETラインがLowの場合、コントローラはディセーブル状態になります。つまり、DORレジスタに0を書き込むだけで、コントローラをディセーブルすることができます。

```
void flpydsk_disable_controller () {
    flpydsk_write_dor (0);
}
```

コントローラーの有効化

コントローラを有効にするには、DORでRESETラインをハイにします。また、FDCをDMAモードで動作させたいので、DORでそのビットも設定する必要があります。

```
void flpydsk_enable_controller () {
    flpydsk_write_dor ( FLPYDSK_DOR_MASK_RESET | FLPYDSK_DOR_MASK_DMA );
}
```

コントローラが無効化された後に有効化されると、割り込みが発生します。この間に、コントローラやドライブの設定を再初期化する必要があります。

FDCの初期化

コントローラのリセット時には、コントローラを再初期化する必要があります。コントローラをリセットすると、IRQ 6が起動します。IRQ 6が発行された後、FDCに接続されているすべてのドライブにSENSE_INTERRUPTコマンドを送信する必要があります(flpydsk_check_intを4回呼び出します)。

その後、コントローラの再設定を行います。CCR(Configuration Control Register)には、データレートを設定するビットが2つしかないことを覚えておいてください。両方を0にすることで、データレートを500Kbpsに設定します。この値はデフォルトの値として最適です。

次に flpydsk_drive_data を呼び出し、Fix Drive Data / Specify コマンドをコントローラに送信して、以下のようなドライブの機械的情報を設定します。ステップレート、ヘッドのロードとアンロードの時間、DMAモードをサポートしているかどうかなど、ドライブの機械的な情報を設定します。

そして、0番シリンダーになるようにドライブを再調整します。

```
void flpydsk_reset () {
    uint32_t st0, cyl;

    //! reset the controller
    flpydsk_disable_controller ();
    flpydsk_enable_controller ();
    flpydsk_wait_irq ();

    //! send CHECK_INT/SENSE INTERRUPT command to all drives
    for (int i=0; i<4; i++)
        flpydsk_check_int (&st0,&cyl);

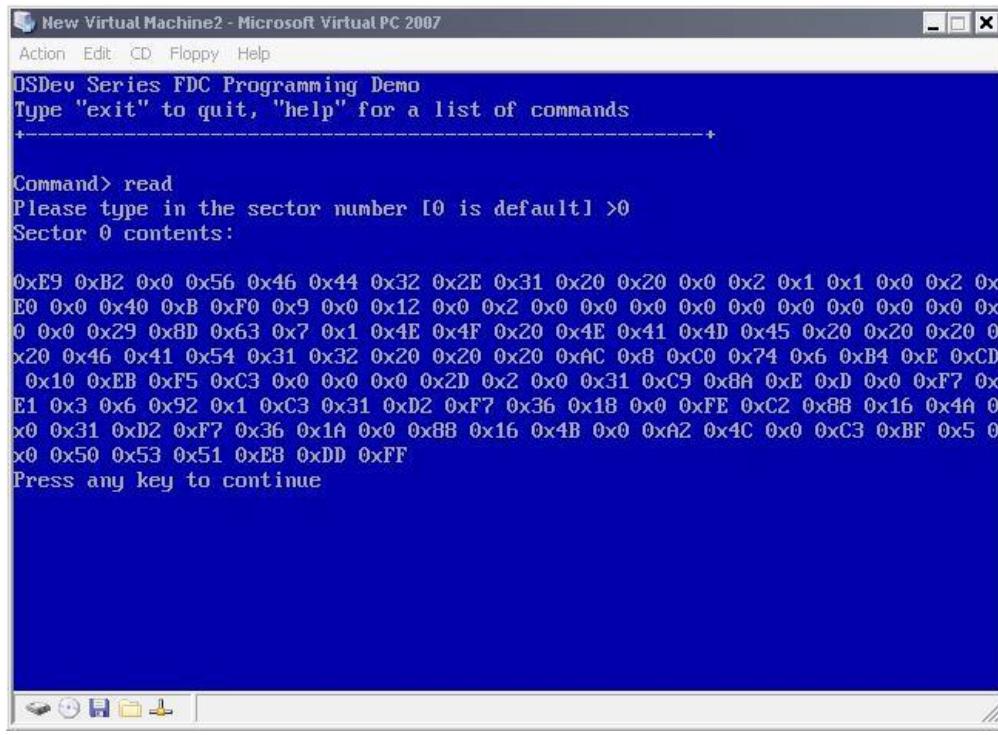
    //! transfer speed 500kb/s
    flpydsk_write_ccr (0);

    //! pass mechanical drive info. steprate=3ms, unload time=240ms, load time=16ms
    flpydsk_drive_data (3,16,240,true);

    //! calibrate the disk
    flpydsk_calibrate ( _CurrentDrive );
}
```

リセットされた後、ドライブは私たちが使用できる状態になります。

Demo



FDCデモの様子

デモダウンロード

注：このデモには、VPCが最初に読み込んだセクターしか読み込まないという既知のバグがあります。これはできるだけ早く解決される予定です。Bochsで実行した場合、既知の問題はありません。

Updates and Changes

文字列から整数への変換 - stdio.h/stdio.cpp

このデモをよりインタラクティブなものにするために、文字列を整数に変換するのに使われる3つの関数を標準ライブラリに入れました。これには、`strtol`, `strtoul`, `atoi`が含まれます。このデモでは`atoi`を使って、ユーザーから入力された文字列を使用可能な整数に変換しています。

フロッピードライバのインストール - flpydsk.cpp

フロッピードライバには、デモで簡単に設定できるように、優れたインストールルーチンが用意されています。HALの`setvect()`ルーチンを使って割り込みハンドラをインストールし、転送用のDMAを初期化し、コントローラをリセットして使用可能な状態にしているだけです。

```
void flpydsk_install (int irq) {
    //! install irq handler
    setvect (irq, i86_flpy_irq);

    //! initialize the DMA for FDC
    flpydsk_initialize_dma ();

    //! reset the fdc
    flpydsk_reset ();

    //! set drive information
    flpydsk_drive_data (13, 1, 0xf, true);
}
```

デモでは、初期化時にこの関数を呼び出し、ドライバーからの読み取りを試みる前にドライバーを設定します。

任意のセクタの読み取り - LBAとCHS - flpydsk.cpp

ドライバは、CHSの詳細を2つの優れた関数の後ろに隠しています。ドライブはCHS(Cylinder/Head/Sector)で動作し、LBA(Linear Block Addressing)については何も知らないことを知っているので、この2つの間で変換するルーチンを提供しなければなりません。これにより、物理的なCHSを気にすることなく、セクタ番号を渡して読み書きすることができます。

LBAをCHSに変換する公式を覚えていませんか？ここではそれを応用してみましょう。

```
void flpydsk_lba_to_chs (int lba,int *head,int *track,int *sector) {
    *head = ( lba % ( FLPY_SECTORS_PER_TRACK * 2 ) ) / ( FLPY_SECTORS_PER_TRACK );
    *track = lba / ( FLPY_SECTORS_PER_TRACK * 2 );
    *sector = lba % FLPY_SECTORS_PER_TRACK + 1;
}
```

2021/11/15 13:10

Operating Systems Development Series

FLPY_SECTORS_PER_TRACKは18です。素晴らしいですね。これで、この関数を呼び出すだけで、任意のリニアセクター番号をCHSの位置に変換できるようになりました。かっこいいでしょう？

ディスクから任意のセクタを読み込めるようにしたいので、そのためのルーチンを用意します。また、すでに

`fipydsk_read_sector_imp`は、読み取りコマンドをコントローラに送信するコードを含んでおり、このルーチンは非常にシンプルです。

```
uint8_t* fipydsk_read_sector (int sectorLBA) {
    if (_CurrentDrive >= 4)
        return 0;

    //! convert LBA sector to CHS
    int head=0, track=0, sector=1;
    fipydsk_lba_to_chs (sectorLBA, &head, &track, &sector);

    //! turn motor on and seek to track
    fipydsk_control_motor (true);
    if (fipydsk_seek (track, head) != 0)
        return 0;

    //! read sector and turn motor off
    fipydsk_read_sector_imp (head, track, sector);
    fipydsk_control_motor (false);

    //! warning: this is a bit hackish
    return (uint8_t*) DMA_BUFFER;
}
```

デモがセクタを読みたいときはいつでも、このルーチンを呼び出します。このルーチンは、セクタをディスク上の物理的な位置(CHS)に変換します。モーターをオンにして、このセクターがあるシリンダーにシークします。その後、`fipydsk_read_sector_imp`を呼び出して、実際にセクタを読み出すマジックを行い、その後、モーターをオフにします。

`fipydsk_read_sector_imp`コールの後、セクタのデータはDMAバッファにあるはずです。このバッファへのポインタを返します。このバッファには、先ほど読み込まれたセクター・データが入っています。かっこいいでしょう？

これは、すべてを結びつける魔法のルーティンです。)

新規読み取りコマンド - main.cpp

このデモは、前回のデモをベースにしています。そのため、前回のデモで構築されたコマンドラインインターフェース (CLI) をそのまま使用しています。これにより、このデモは今までで最も複雑なデモにもなっています。

CLIのコマンドリストに新しいコマンド「read」を追加しました。このコマンドでは、ディスクから任意のセクタを読み取ることができます。これは、このチュートリアルで作成されたフロッピードライバを使用して行います。

このコマンドは関数の中にあり、デモではreadと入力して実行します。これは512バイトを4つの128バイトのブロックに分けて読みやすくダンプします。各ブロックが終わると、次のチャンクに進むためにキーを押すように促されます。新しいatoi関数を使って、入力されたセクタ番号 (LBAセクタ番号) をintに変換し、それを読み込みます。読者の皆さん、これが魔法を起こす関数です。

```
void cmd_read_sect () {
    uint32_t sectornum = 0;
    char sectornumbuf [4];
    uint8_t* sector = 0;

    DebugPrintf ("\n\rPlease type in the sector number [0 is default] >");
    get_cmd (sectornumbuf, 3);
    sectornum = atoi (sectornumbuf);

    DebugPrintf ("\n\rSector %i contents:\n\n\r", sectornum);

    //! read sector from disk
    sector = fipydsk_read_sector ( sectornum );

    //! display sector
    if (sector!=0) {

        int i = 0;
        for ( int c = 0; c < 4; c++ ) {

            for (int j = 0; j < 128; j++)
                DebugPrintf ("0x%02x ", sector[ i + j ]);
            i += 128;

            DebugPrintf ("\n\rPress any key to continue\n\r");
            getch ();
        }
    }
    else
        DebugPrintf ("\n\r*** Error reading sector from disk");

    DebugPrintf ("\n\rDone.");
}
```

結論

長いチュートリアルになってしまいましたね。改良のためにいくつか変更を加えて、より良い完成度の高いものにしていくかもしれません。;)

次回のチュートリアルでは、DMAについて説明します。DMAをプログラミングするためのインターフェイスを作成し、FDCドライバの中でそれをうまく利用していきます。この後は、またファイルシステムの話ですね。心配しないでください - この後はマルチタスクです!

次の機会まで。

~マイク

BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中です。質問やコメントはありますか？お気軽にお問い合わせください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。



第19章 ホーム

Chapter 19

[Home](#)



オペレーティングシステム開発シリーズ

オペレーティング・システムの開発 - 8237A ISA DMAC

by Mike, 2009

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

注：今後、デモ名は「Demo00」（00は章名）という形式で表記します。これは、デモの名前と章の名前が関連しないという問題を解決するためと、読者が特定のデモがどの章を参照しているかを簡単に知ることができるようにするためです。古いチャプターもこの設定で更新されます。すべてのチャプターが更新されたら、このコメントは削除されます。

また、Virtual PCのバグは、本章では修正されていますが、前章ではまだ修正されていません。前の章のデモはVirtual PCでうまく動作しませんでしたが、この章のデモはDMAコードのマイナーなバグフィックスとアップデートでうまく動作するようになりました。このデモはVirtual PCとBochsでテストされ、動作しました。前の章がこの修正で更新されたら、このコメントは削除されます。

Introduction

歓迎します。:)

この章では、DMAC (Direct Memory Access Controller) について詳しくご紹介します。DMACは、ソフトウェアを使わずに、デバイスからのデータブロックを直接メモリに転送する方法を提供します。これにより、ソフトウェアではなくハードウェアがデータを転送するため、非常に高速なデータ転送が可能になります。

今日のリストはこちらです。

DMAの歴史

DMAハードウェ

ア DMAポート

DMAレジスター

DMAコマンド

アブストラクト

ダイレクトメモリアクセス (DMA) とは、最近のコンピュータに搭載されている機能で、プロセッサとのやりとりなしに大きなデータブロックを移動させることができるデバイスです。この機能は、フロッピーディスクのプログラミングの章でご紹介したように、非常に便利です。デバイスがデータブロックを転送している間、プロセッサはデータがメモリや他のデバイスに転送されているかどうかを気にすることなく、ソフトウェアの実行を続けることができます。基本的な考え方は、DMAデバイスが自分でタスクを実行するようにスケジュールすることです。かっこいいでしょう？

バスやアーキテクチャの設計によって、ダイレクトメモリアクセスを実行する方法は異なります。今回はISA Direct Memory Access Controller (DMAC)に焦点を当てますが、念のため他の方式についても取り上げることにしました。

ISA

ISA (Industry Standard Architecture) では、インテル8237マイクロコントローラをベースにしたコントローラを介して、DMAリクエストを集中的に行うことができます。ATXマザーボードの設計では、コントローラは1つしかありませんでした。しかし、ATX以降のアーキテクチャでは、コントローラは2つあります。これは、PIC (Programmable Interrupt Controllers) がスレーブ化されているのと同じです。両方のコントローラは常に4MHzで動作します。

性能やデバイスの数が限られているため、新しいデバイスではPIOやUDMAを代わりに使う傾向があります。しかし、レガシーデバイスではISAでDMAがサポートされています。

これらのデバイスはすべて、コントローラのチャンネルに接続されている。これらのチャンネルとともに、各チャンネルにはDACK (DMA Acknowledge) ラインとDRQ (DMA Request) があります。

ここでは、両DMAC (Direct Memory Access Controllers) の標準的な割り当てを紹介します。

- XT:

チャンネル0: システムで使用、使用不可 (DRAM Refresh、廃止予定)

チャンネル1：使用可能、標準的なDMA割り当てなし

チャンネル2：フロッピーディスクコントローラ

チャンネル3：ハードディスク・コントローラ（PIOまたはUDMAを推奨） ATのみ。

チャンネル4：XTコントローラにカスケード接続 - スレーブDMAコントローラをマスターに入力

チャンネル5：使用可能、標準的なDMA割り当てなし（16ビット）

チャンネル6：使用可能、標準的なDMA割り当てなし（16ビット）

チャンネル7：使用可能、標準的なDMA割り当てなし（16ビット）

転送を開始するために、ソフトウェアはチャネルアドレスとカウントレジスタに、物理メモリ内の転送を完了する場所と転送サイズを設定します。その後、そのメモリブロックからの読み出しありは書き込みを設定し、コントローラを転送完了に向けて動作させます。転送が完了すると、転送を開始したデバイスからIRQ（Interrupt Request）が発行され、システムソフトウェアがそれをキャッチして処理を行います。これは重要なことです。以上が、DMAを使って転送を開始する際に必要となる手順です。

PCI

PCIデバイスは同じDMAコントローラを共有しておらず、中央にDMAコントローラがあるわけでもない。代わりに、PCIローカルバス上のPCIデバイスが、PCIバスコントローラーにバスマスターになることを要求する（バスをコントロールする）。その後、物理メモリへの読み書きのリクエストがノースブリッジに渡され、ノースブリッジはそのリクエストをメモリ操作に変換してメモリコントローラに送ります。

PCIの転送は4GBの物理メモリに制限されています。しかし、デバイスとPCIブリッジがDouble Address Cycle (DAC)または同様の技術を実装している場合、PCIコントローラが4GBの物理メモリを超えて読み書きのリクエストを開始することができます。

ISA DMAハードウェア

ダイレクトメモリアクセスコントローラ (DMAC)

ISA (Industry Standard Architecture) では、オリジナルのIntel 8237 DMAチップをベースにしたコントローラを使用しています。最新のDMACはより多くの機能を備えていますが、ほとんどが8237マイクロコントローラとの互換性を持っています。新しいPCにはより高度なDMACが搭載されていますが、すべての始まりとなったデバイスを見ることはいつでも素晴らしいことです。そこで、デュアル・インライン・パッケージ (DIP) で配布されているオリジナルの8237Aコントローラのピンダイアグラムをご紹介します。

IOR	1	40	A7
IOW	2	39	A6
MEMR	3	38	A5
MEMW	4	37	A4
*	5	36	EOP
READY	6	35	A3
HACK	7	34	A2
ADSTB	8	33	A1
AEN	9	8237A	32
HREQ	10	DMAC	31
CS	11	30	DB0
CLK	12	29	DB1
RESET	13	28	DB2
DACK2	14	27	DB3
DACK3	15	26	DB4
DREQ3	16	25	DACK0
DREQ2	17	24	DACK1
DREQ1	18	23	DB5
DREQ0	19	22	DB6
GND/Vss	20	21	DB7

これが、この章でプログラミングすることになるコントローラです。たくさんのピンがありますが、それほど複雑ではありません。重要なものを中心に見ていきましょう。

Pin 1 (IOR) I/O Read

Pin 2 (IOW) I/O Write

Pin 3 (MEMR) メモリリード Pin 4

(MEMW) メモリライト Pin 5

ピン6 (READY)

Pin 7 (HACK) Hold Acknowledge Pin

8 (ADSTB) Address Strobe Pin 9

(AEN) Address Enable

Pin 10 (HREQ) ホールドリクエスト

Pin 11 (CS) チップセレクト

Pin 12 (CLK) クロック

端子13 (RESET) リセット

ピン14-15 (DACK) DMA アカウレッジ

ピン16-19 (DREQ0-DREQ3) DMAリクエスト

Pin20 (GND/Vss) グラウンド

ピン21-23 (DB0-DB3) データバス ピン24-

25 (DACK) DMA Acknowledge ピン26-30

(DB4-DB7) データバス ピン31 (Vcc) +5V

電源

32-35番ピン (A0-A3) アドレスライン

Pin 36 (EOP) End Of Process

Pins 37-40 (A4-A7) アドレスライン

これは悪くないですね。20番ピンがグランド、31番ピンが電源です。12番ピン (CLK) は、どのコントローラにも共通しているものです。プロセッサのCLK端子に接続し、コントローラ内の動作タイミングを制御するクロック信号を入力します。CS (チップセレクト) 端子も、ほとんどのコントローラに共通して見られるものだ。データバス上のI/Oデバイスとしてコントローラを選択するために使用されます。

RESETは、コントローラの内部レジスタ (Status,

リクエスト、テンポラリー、コマンド）、内部のフリップフロップをクリアし、マスクレジスタを設定します。ここまででは、特に目新しいことはありませんね。システムアドレスバスに接続するゲルネリックアドレスラインA0～A7があります。入力時には、CPUはA0-A3にデータを書き込み、読み出すレジスタを選択することしかできません。すべてのピンは（物理的なメモリアドレスへの）出力に使用されますが、DMAリクエスト時にのみ有効になります。最後に、システムデータバスに接続する汎用のD0-D7ピンです。

次に、より興味深いピンについて説明します。ここまでで説明したように、DMACはシステムのアドレスとデータバスに接続していることがわかりました。読者の多くは、このことにあまり驚かないでしょう。しかし、ご想像のとおり、DMACはCPUからの直接の注意を必要とします。そのため、DMACがCPUと通信できるように、CPUに接続するラインをいくつか用意しており、その逆も可能です。これは、HACKおよびHREQピンで行われます。HACK (Hold Acknowledge) は、CPUがDMACにシステムバスの完全な制御権を与えたときにHighに保たれます。これによりDMACは、メモリコントローラにデータを送信しても大丈夫なタイミングを知ることができます。結局のところ、DMACとプロセッサーの両方が同じシステムバスを同時に使用することはできないのだ。

DMACは、システムバスがプロセッサーによって使用されていない場合に限り、物理メモリに直接データを転送する。DMACがシステムバスを必要とするのは、メモリコントローラにデータを送信し、メモリ変換や物理メモリへの読み書きを行うためである。

なるほど、CPUはDMACにシステムバスの乗っ取りを伝える方法を持っているんですね。しかし、そもそもDMACはどうやってCPUにシステムバスの必要性を伝えるのだろうか？それが、HREQ (Hold Request) ラインである。DMAに接続されたデバイス（フロッピーコントローラなど）からDMAリクエスト (DRQ) がトリガーされ、その「チャンネル」が現在無効になっていない場合、コントローラーは次のクロックサイクルでHREQをハイレベルにして、リクエストを完了するためにシステムバスの制御が必要であることをCPUに通知する。

DR0～DR3 (DMAリクエストライン) は、デバイスがDMAにリクエストを通知するために使用するラインです。例えば、フロッピー・ドライブ・コントローラ (FDC) は、通常、DR2 (「チャンネル2」) を使用するように接続されています。そこで、そのチャンネルを有効にして、FDCがDMACを使用するようにプログラムすると、FDCにリードまたはライトコマンドが送られると、FDCはRQ2ラインをアクティブにして、DMACに注意が必要であることを通知する。ここからは、そのチャンネルをプログラムしたモードに応じて、DMACを介してすべての読み取りまたは書き込みが行われます。

ここで、もうひとつ重要なポイントがあります。DRQラインが4本しかないことを知っていること。1台のDMACに接続できる機器は4台まで。これはかなり限定的ですよね。i86アーキテクチャでは、2つのDMACをくっつけることで、この問題をある程度解決しています。ここでは、その方法を紹介します。

これまでのところ、すべてが順調に進んでいます。ソフトウェアがCPUにDMACのプログラムを指示することはわかったが、CPUはDMACにレジスタの読み書きが必要であることをどのように伝えるのだろうか？それがIOR (I/O Read) とIOW (I/O Write) のピンです。同様に、DMACはメモリコントローラに対して、MEMR (Memory Read) やMEMW (Memory Write) を出力して、メモリの読み出しや書き込みの制御線をアクティブにすることで、読み出しや書き込みを行なうことを伝えている。EOP (End Of Process) は、リクエストが完了したときにデバイスに信号を送るためのものです。リクエストが完了するのは、そのチャンネルのターミナルカウント (TC) に到達したときです。これはプログラム可能なカウンター値です。AEN (Address Enable) は、コントローラが内部の8ビットアドレスラッチレジスタをシステムアドレスバスにロードすることを通知するために使用されます。ADSTB(Address Strobe)は、上位アドレスバイトを外部ラッチレジスタにストロープするために使用します。

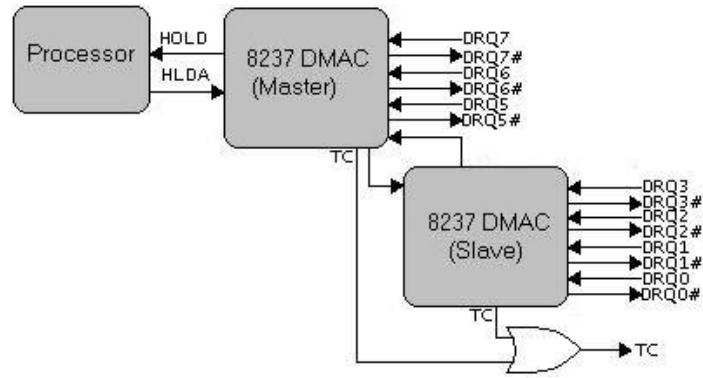
いろいろあるんですね。コントローラが行う操作の詳細は、モードや転送タイプによって異なります。しかし、基本的な手順は同じです。デバイスがDMACに通知し、DMACがシステムバスを介してCPUに制御を通知する。DMACは制御を待ちます。DMACは制御を待ち、制御を受けるとチャンネル・アドレス・レジスタを内部ラッチ・レジスタにロードします。そこから、MEMR、MEMWを設定し、必要に応じてメモリの読み書きを行うことになる。待てよ、何だって？メモリからの読み出しがわかると思いますが、書き込みをしたらどこに行くのでしょうか？

もう一度FDCの章を見てみましょう。DMACやCPU、メモリコントローラなどと同じデータバスに接続するピンD0～D7があることに注目してください。つまり、メモリからの書き込み時に必要なのは、MEMRラインをアクティブにしてアドレスバスにアドレスをアップロードし、メモリコントローラがデータを変換してデータバスに載せることだけなのだ。FDCは書き込み要求を待っているので、読み込んだデータを取り込み、FDCに送られてきた書き込みコマンドで設定されたディスクに書き込む。ディスクからの読み出しの場合も基本的には同じですが、DMACは代わりにMEMWラインをアクティブにします。メモリコントローラーは、FDCから送られてきたデータバスから書き込むべきデータを取り込みます。すべてが順調に進むと、DMACはプロセッサーのHREQラインを解放し、CPUが再びバスを完全に制御できるようになります。

ここで重要なのは、プロセッサーはDMACの終了を待つことはできないということです。CPUは、システムバスへのアクセスが必要になると、HACKラインをローレベルにします。その間、DMACは処理を続けるためにHACKラインが再びハイレベルになるまで待つしかありません。読者の皆さん、こんにちは。i86では、DMACをもう1つ追加して、使用できるチャンネル数を8にしました。まあ、そんな感じです。それでは見てみましょう。

x86のDMAC

最近のPCには2つのDMACが搭載されているのを覚えていますか？両方のDMACは、2つのPICが接続されているのと同じような方法で接続されています…ただ、逆方向です。ええっ！？そうなんですよ。）では、見てみましょう。



DMACは、ISAバスの制御を行う際に、プロセッサのHOLDおよびHLDA (Hold Acknowledge) ピンを使用する。DMACはHOLDを通じてプロセッサに信号を送り、プロセッサはHLDAを通じてこの要求を承認する。また、プライマリDMACがDRQの5~7を持っているのに対し、セカンド（スレーブ）にはDRQの0~3が含まれていることにも注目してほしい。DRQとは、DMAリクエストのことである。これらのラインは、DMACを使用するシステム内のさまざまなデバイスに接続する。デバイスがDMACの使用を要求すると、ラインをハイレベルにしてDMACに信号を送ります。画像をもう一度見てみると、面白いことがわかるかもしれません。DRQ4はどこにある？

DRQ4は両方のデバイスに存在しますが、それぞれのDMACを接続するものです。それらは画像に示されています（ラベルは付いていません）。DRQラインはDMACに信号を送るために使われる所以、これによってプライマリDMACとセカンドDMACがお互いに正しいDRQラインを上げるように信号を送ることができます。つまり、DMACをプログラミングする際には、DRQ4がプライマリコントローラとスレーブコントローラを接続するために使用されていることを忘れてはならないのです。そのため、DRQ4は使用できません。上の図を振り返ると、プライマリDMACまたはスレーブDMACのいずれかが完了した（TC (Terminal Count) ラインを上げた）場合にtrueを出力するORゲートも見られます。TCラインは、DMACに送信された転送要求が完了したときに上がります。

さて、ここで覚えておかなければならぬ重要なことを、参考のためにまとめておきましょう。

DMAは常に物理メモリ上で動作し、仮想メモリ上では動作しない

i86アーキテクチャのDMACを使用するには、8台のデバイスしか接続できません。

DRQ4（チャンネル4）は、プライマリDMACとセカンドDMACの接続に使用されており、使用することはできません。

また、これらがどのように構成されているかについても興味深いものがあります。スレーブDMACは、マスターDMACに接続する最初のDMACであり、その逆ではないのです。スレーブDMACがチャンネル0~3（厳密にはプライマリDMACへの接続に使われる4も）を担当し、プライマリDMACがチャンネル5~7を担当するのはこのためです。なんだか変ですよね？このように、2つのPICが一緒に動作する方法とは多少異なります。また、これらのコントローラーが接続されているため、マスターDMACは16ビットDMACのように動作し、スレーブDMACは8ビットDMACのように動作することにも注意が必要です。これが原因です。

1台目のDMACがスレーブ（8ビット）、2台目がマスター（16ビット）。

ISA DMAインターフェース

ポートマッピング

DMAコントローラーが2つあるので、ポートも2組あります。

汎用レジスター

ISA DMA Ports		
DMAC 0 Port (Slave)	DMAC 1 Port (Master)	Description
0x08	0xD0	Status Register (Read)
0x08	0xD0	Command Register (Write)
0x09	0xD2	Request Register (Write)
0x0A	0xD4	Single Mask Register (Write)
0x0B	0xD6	Mode Register (Write)
0x0C	0xD8	Clear Byte Pointer Flip-Flop (Write)
0x0D	0xDA	Intermediate Register (Read)
0x0D	0xDA	Master Clear (Write)
0x0E	0xDC	Clear Mask Register (Write)
0x0F	0xDE	Write Mask Register (Write)

これらのレジスターについては、次のセクションで詳しく説明します。これらのレジスターは、両方のDMACを操作する際に使用され

ます。これらのレジスタは、ポートマップドI/Oを通して読み書きすることができます。つまり、標準的なi86のinおよびout命令を使用します。

DMACは逆向きであることを覚えておくことはとても重要です。DMAC 0はスレーブで、DMAC 1はマスターです。また、ポートの範囲が異なることにも注意してください。スレーブは8ビット、マスターは16ビットであることを覚えていますか？読みやすくするために、これらの醜い数字を列举して抽象化してみましょう。

```
enum DMA0_IO {
    DMA0_STATUS_REG = 0x08,
    DMA0_COMMAND_REG = 0x08,
    DMA0_REQUEST_REG = 0x09,
    DMA0_CHANMASK_REG = 0xa,
    DMA0_MODE_REG = 0xb,
    DMA0_CLEARBYTE_FLIPFLOP_REG = 0xc,
    DMA0_TEMP_REG = 0xd,
    DMA0_MASTER_CLEAR_REG = 0xd,
    DMA0_CLEAR_MASK_REG = 0xe,
    DMA0_MASK_REG = 0xf
};
```

これらの値が上の表と一致していることに注目してください。次にDMAC 2について。

```
enum DMA1_IO {
    DMA1_STATUS_REG = 0xd0,
    DMA1_COMMAND_REG = 0xd0,
    DMA1_REQUEST_REG = 0xd2,
    DMA1_CHANMASK_REG = 0xd4,
    DMA1_MODE_REG = 0xd6,
    DMA1_CLEARBYTE_FLIPFLOP_REG = 0xd8,
    DMA1_INTER_REG = 0xda,
    DMA1_UNMASK_ALL_REG = 0xdc,
    DMA1_MASK_REG = 0xde
};
```

それでは、レジスターをご覧ください。

チャンネルレジスター

i86では、上記のレジスタに加えて、各チャンネルのアドレスやカウンターを制御するための以下のレジスタが用意されています。

ISA DMAC Channel Ports		
DMAC 0 Port (Slave)	DMAC 1 Port (Master)	Description
0x0	0xC0	Channel 0 Address/Channel 4 Address
0x1	0xC2	Channel 0 Counter/Channel 4 Counter
0x2	0xC4	Channel 1 Address/Channel 5 Address
0x3	0xC6	Channel 1 Counter/Channel 5 Counter
0x4	0xC8	Channel 2 Address/Channel 6 Address
0x5	0xCA	Channel 2 Counter/Channel 6 Counter
0x6	0xCC	Channel 3 Address/Channel 7 Address
0x7	0xCE	Channel 3 Counter/Channel 7 Counter

上の表をもう一度見てください。マスターDMACのチャンネル0のアドレスは.....えっ、ポート0！？このシリーズでは、i/oポート0を見つけてことで、歴史的な瞬間を迎えます。）

また、マスターDMACが16ビットであるのに対し、スレーブDMACは8ビットであることも覚えているだろうか。これは重要な特徴で、特にここでは、スレーブDMACには8ビットの値を、マスターDMACには16ビットの値を読み書きできることを意味しています。

さて、これらのレジスタの詳細を説明する前に、まずこれらを隠すことにしましょう。下の例を見ると、派手なことは何も行われていないことがわかるでしょう。これらのレジスタはすべてポートマップドI/Oでアクセスされることを覚えておいてください。言い換えれば、x86マシン命令のin/outを使って読み書きすることができます。

```
enum DMA0_CHANNEL_IO {
    DMA0_CHAN0_ADDR_REG = 0, //! That's right, i/o port 0
    DMA0_CHAN0_COUNT_REG = 1,
    DMA0_CHAN1_ADDR_REG = 2,
    DMA0_CHAN1_COUNT_REG = 3,
    DMA0_CHAN2_ADDR_REG = 4,
    DMA0_CHAN2_COUNT_REG = 5,
    DMA0_CHAN3_ADDR_REG = 6,
    DMA0_CHAN3_COUNT_REG = 7,
};
```

...そしていよいよDMAC2へ...。

```
enum DMA1_CHANNEL_IO {
    DMA1_CHAN4_ADDR_REG = 0xc0,
    DMA1_CHAN4_COUNT_REG = 0xc2,
    DMA1_CHAN5_ADDR_REG = 0xc4,
    DMA1_CHAN5_COUNT_REG = 0xc6,
    DMA1_CHAN6_ADDR_REG = 0xc8,
    DMA1_CHAN6_COUNT_REG = 0xca,
    DMA1_CHAN7_ADDR_REG = 0xcc,
    DMA1_CHAN7_COUNT_REG = 0xce,
};
```

これらのレジスタの基本的な目的は、チャンネルをどのように開始するかをDMACに伝える方法を提供することです。各チャンネルには、ベースアドレスとカウンターがあります。ベースアドレスは、読み書きを開始するメモリ上の位置で、カウンターは、そのチャンネルでどれだけ転送するかをDMACに伝えます。ここで重要なのは、これらが仮想ではなく常に物理的なアドレスであるということです。

例を挙げてみましょう。チャンネルが使用するベースアドレスを設定するには、上記の表に示されている正しいi/oポートに書き込むだけです。DMA0_CHAN0_ADDR_REGが0で、DMA1_CHAN7_ADDR_REGが表の最後の値(0xde)だとすると、これは簡単です。すべてのサンプルコードは、この章の最後にあるデモに含まれています。

```
void dma_set_address(uint8_t channel, uint8_t low, uint8_t high)
{
    if (channel > 8)
        return;

    unsigned short port = 0;
    switch (channel) {
        case 0: {port = DMA0_CHAN0_ADDR_REG; break;}
        case 1: {port = DMA0_CHAN1_ADDR_REG; break;}
        case 2: {port = DMA0_CHAN2_ADDR_REG; break;}
        case 3: {port = DMA0_CHAN3_ADDR_REG; break;}
        case 4: {port = DMA1_CHAN4_ADDR_REG; break;}
        case 5: {port = DMA1_CHAN5_ADDR_REG; break;}
        case 6: {port = DMA1_CHAN6_ADDR_REG; break;}
        case 7: {port = DMA1_CHAN7_ADDR_REG; break;}
    }

    outportb(port, low);
    outportb(port, high);
}
```

よく似た方法で、特定のチャンネルのカウントレジスタを設定するルーチンを書くことができます。

```
void dma_set_count(uint8_t channel, uint8_t low, uint8_t high)
{
    if (channel > 8)
        return;

    unsigned short port = 0;
    switch (channel) {
        case 0: {port = DMA0_CHAN0_COUNT_REG; break;}
        case 1: {port = DMA0_CHAN1_COUNT_REG; break;}
        case 2: {port = DMA0_CHAN2_COUNT_REG; break;}
        case 3: {port = DMA0_CHAN3_COUNT_REG; break;}
        case 4: {port = DMA1_CHAN4_COUNT_REG; break;}
        case 5: {port = DMA1_CHAN5_COUNT_REG; break;}
        case 6: {port = DMA1_CHAN6_COUNT_REG; break;}
        case 7: {port = DMA1_CHAN7_COUNT_REG; break;}
    }

    outportb(port, low);
    outportb(port, high);
}
```

ここで重要なのは、これらのレジスターが16ビットであるということです。これは、DMACが1つのチャンネルから一度に転送できるのは最大でも64kであることを意味しています。

また、これらは物理的なアドレスであることにも注意が必要です。システムソフトウェアがページングを有効にしている場合、チャネルが使用する場所を、使用されるメモリの領域を識別して同じ仮想アドレスにマッピングする必要があります。

そこで、おさらいです。8つのチャネルがあることを知り、そのチャネルのデバイスがDMACを使えるようにした後、チャネルレジスタの1つに書き込むことで、チャネル情報（メモリの位置、読み書きの条件）を与えて、DMACへの読み書きを開始することができる。このデータはどこから来るのかと聞かれるかもしれません。また、読み出した場合、そのデータはどこに行くのか？これは、そのチャネルを制御しているデバイス次第です。例えば、フロッピーディスクドライブでは、リードコマンドをフロッピーディスクドライブコントローラー（FDC）に送ると、FDCはDMACに通知して転送を開始します。DMACは、ベースとなる物理アドレス、チャネルの動作（リードかライトか、この場合はリードであることが望ましい）、バッファのサイズを取得し、あとは自分で書き込みます。FDCはDMACにデータを転送し続け、DMACはそのチャネルに格納されているアドレスが指すバッファにデータを配置することになる。FDCはDMACにデータを転送し続け、DMACはそのチャネルに格納されているアドレスが指すバッファにデータを入れます。ここでは、そのチャネルのアドレスとカウントレジスタに書き込むことで、バッファの位置とサイズを設定します。

待って、待って、待って。DMACは一度に64Kしか転送できないことを覚えていますか？さらに悪いことに 各チャネルのベースアドレスも同じ制限があり、DMACがアクセスできるのは64KのRAMに限られるということになります。これは悪い制限だと思いませんか？これを解決するのが、外部ページレジスターです。もっと詳しく見てみましょう。

拡張ページアドレスレジスター

ページレジスタは、チャネルが設定されているメモリロケーションがどのページに存在するかを設定するために使用されます。この8ビットをチャネルのベースアドレスに追加すると（チャネルのベースアドレスを0xFFFFFFFとする）、実質的に8ビット増えることになり、最大16MBのメモリにアクセスできるようになります。これがページレジスターの仕組みです。

これらのページレジスターには、そのチャネルの転送アドレスの上位8ビットのみが格納されます。これは、これらのページレジスターの値が常に64kの倍数であることを意味し、重要な特徴です。

案の定、ここでちょっとした問題が発生する。DMACが1つだった当初のパソコンは、AT/EISA/MCAとそれ以降のパソコンとではi/oポートが異なっていた。また、DMACが2つになったことで、レジスターが追加され、ビット数も増えた。オリジナルのPCのページレジスターは4ビットしか追加されていない（A16～A19）。一方、新しいコンピュータでは、ベースチャネルアドレスに8ビット（A16～A23）が追加された。

ISA DMAC Extended Page Address Registers	
Port	Description
0x80	Channel 0 (Original PC) / Extra / Diagnostic port
0x81	Channel 1 (Original PC) / Channel 2 (AT)
0x82	Channel 2 (Original PC) / Channel 3 (AT)
0x83	Channel 3 (Original PC) / Channel 1 (AT)
0x84	Extra
0x85	Extra
0x86	Extra
0x87	Channel 0 (AT)
0x88	Extra
0x89	Channel 6 (AT)
0x8A	Channel 7 (AT)
0x8B	Channel 5 (AT)
0x8C	Extra
0x8D	Extra
0x8E	Extra
0x8F	Channel 4 (AT) / Memory refresh / Slave Connect

さて、ここでちょっと立ち止まってみましょう。*さて、上の表で気にしなければならないのは、ATのポートだけです。これは、すべてのチャネル外部ページレジスタが、チャネルの設定時に保存したチャネルのベースアドレスにさらに8ビットを追加することを意味します（前のセクションを参照）。例えば、フロッピーコントローラをプログラミングする際、フロッピーがDMAチャネル2を使用することがわかっています。例えば、フロッピーディスクコントローラをプログラムする際、フロッピーディスクはDMAチャネル2を使用することがわかっています。例えば、64kよりも低い位置にバッファを保存したい場合、チャネルのアドレスをどこかに設定すればよいのです。まあ、そんなところです。また、そのアドレスの上位8ビットを決定するために使用されるページレジスタを設定する必要があります。つまり、ページレジスタを設定するには

0に設定すると、ページ0、アドレスに何も追加されませ

ん1に設定すると、ページ1、アドレスに64kが追加され

ます

2に設定すると、ページ2、128Kがアドレスに追加される

255に設定すると、Page 255 = $255 \times 64\text{K} = 0xFF0000$ となり、上位8ビット全てが設定され、16、320K、約16MBがアドレスに追加される。

ページテーブルのページを変更すると、DMAが読み書きするアドレスが変わることに注目してほしい。これにより、DMACは最大16MBのメモリーに効率よくアクセスできるようになります。すごいでしょう？まだ少し制限がありますが、64Kに制限されるよりはずっと良いと思いませんか？

他のレジスターと同様に、醜いマジックナンバーを隠すことができます。

```
enum DMA0_PAGE_REG {
    DMA_PAGE_EXTRA0 = 0x80, //! Also diagnostics port
    DMA_PAGE_CHAN2_ADDRBYTE2 = 0x81,
    DMA_PAGE_CHAN3_ADDRBYTE2 = 0x82,
    DMA_PAGE_CHAN1_ADDRBYTE2 = 0x83,
    DMA_PAGE_EXTRA1 = 0x84,
    DMA_PAGE_EXTRA2 = 0x85,
    DMA_PAGE_EXTRA3 = 0x86,
    DMA_PAGE_CHAN6_ADDRBYTE2 = 0x87,
    DMA_PAGE_CHAN7_ADDRBYTE2 = 0x88,
    DMA_PAGE_CHAN5_ADDRBYTE2 = 0x89,
    DMA_PAGE_EXTRA4 = 0x8c,
    DMA_PAGE_EXTRA5 = 0x8d,
    DMA_PAGE_EXTRA6 = 0x8e,
    DMA_PAGE_DRAM_REFRESH = 0x8f //!no longer used in new PCs
};
```

これらのレジスタを設定するために必要なことは、どのレジスタに書き込まれているかを（どのチャンネルが渡されているかで）判断し、そのレジスタに値を書き込むことです。

```
void dma_set_external_page_register (uint8_t reg, uint8_t val)
{
    if (reg > 14)
        return;

    unsigned short port = 0;
    switch (reg) {

        case 1: {port = DMA_PAGE_CHAN1_ADDRBYTE2; break;}
        case 2: {port = DMA_PAGE_CHAN2_ADDRBYTE2; break;}
        case 3: {port = DMA_PAGE_CHAN3_ADDRBYTE2; break;}
        case 4: {return;}//! nothing should ever write to register 4
        case 5: {port = DMA_PAGE_CHAN5_ADDRBYTE2; break;}
        case 6: {port = DMA_PAGE_CHAN6_ADDRBYTE2; break;}
        case 7: {port = DMA_PAGE_CHAN7_ADDRBYTE2; break;}
    }

    outportb(port, val);
}
```

注目すべきは、ケース4がコメントされていることです。チャンネル4はマスターDMACとのカスケードに使われることを覚えていませんか？これがると、何も使えないからです。上記の各ケースは、ページを設定するチャンネルを表しています。つまり、`dma_set_external_page_register (2, 0x1000);`のような呼び出しで、チャンネル2のページレジスタに0x1000を設定することができます。いいですか？

レジスター

上記のレジスタに加えて、コントローラでは以下のレジスタも利用可能です。

コマンドレジスタ

このレジスタは、DMACの制御に使用されます。以下のようなフォーマットになっています。

ビット0 : MMT Memory to Memory Transfer

0 : Disable

1: 有効

ビット1 : ADHEチャンネル0アドレスホールド

0 : Disable

1: 有効

ビット2 : COND Controller Enable

0 : Disable

1: 有効

ビット3 : COMPタイミング

0 : ノーマル

1: コンプレッサー

ビット4 : PRIOプライオリティ
0 : 固定プライオリティ
1 : ノーマルプライオリティ
ビット5 : EXTWライト選択
0: レイトライト選択
1: 拡張ライトセレクション
ビット6 : DROP DMA Request (DREQ)

0: DREQセンス・アクティブ・ハイ
 1: DREQセンス・アクティブ・ロー
 ビット7:DACKP DMAアクノレッジ(DACK)
 0: DACKセンス・アクティブ・ロー
 1: DACKセンス・アクティブ・ハイ

これらのビットのほとんどは、i86アーキテクチャでは動作しません。唯一動作するのはビット2で、これを使ってコントローラを有効にしたり無効にしたりすることができます。メモリからメモリへの直接の転送も便利だと思うでしょうが、そうではありません。他のビットを使うと、何もできなかったり、予想外の結果になったりします。

これらは、本章最後のデモにあるdma.hというヘッダーファイルに含まれています。ここでは、ビットマスクとして表示しています。

```
enum DMA_CMD_REG_MASK {
    DMA_CMD_MASK_MEMTOMEM = 1,
    DMA_CMD_MASK_CHAN0ADDRHOLD = 2,
    DMA_CMD_MASK_ENABLE = 4,
    DMA_CMD_MASK_TIMING = 8,
    DMA_CMD_MASK_PRIORITY = 0x10,
    DMA_CMD_MASK_WRITESEL = 0x20,
    DMA_CMD_MASK_DREQ = 0x40,
    DMA_CMD_MASK_DACK = 0x80
};
```

モードレジスタ（ライト）

このモードは、コントローラのモードを設定します。以下のようなフォーマットになっています。

ビット0~1: SEL0, SEL1 チャンネルセレクト

- 00: チャンネル0
- 01: チャンネル1
- 10: チャンネル2
- 11: チャンネル3

ビット2-3: TRA0, TRA1 転送タイプ

- 00: コントローラのセルフテスト
- 01: ライト転送
- 10: リード転送
- 11: 無効

ビット4: AUTO 転送完了後の自動再初期化（デバイスがサポートしている必要があります。）

ビット5: IDEC

6-7ビット: MOD0, MOD1 モード

- 00: トランスマスター・オン・デマンド
- 01: シングルDMA
- 10: ブロックDMA転送
- 11: カスケードモード

このレジスタは重要です。チャンネルを設定し、メモリブロックを読み書きする準備をするためには、このレジスタに動作モードを書き込む必要があります。しかし、このレジスタに書き込む前に、何かを変更する前に、モードを設定したいチャンネルをマスクオフ（ディセーブル）することをお勧めします。これは、現在使用中のチャンネルのモードを変更すると、データが破損するなどの問題が発生するためです。

まず最初にしたいことは、醜い数字を意味のある名前で隠すことですが、それがここにあります。しかし、これは少し違います。これらの列挙は、マスクとフラグの組み合わせです。マスクは上のリストのビットフォーマットと一致しています。フラグは単純化のためにあります。フラグは、上のリストの必要なビットをセットまたはクリアすることで、オプションをビットワイズオアすることができます。例えば、チャンネル番号とチャンネルのモードを組み合わせて、自動初期化による単一転送の読み取りに設定するには、次のようにします。

| dma_mode_transfer_singleです。かっこいいでしょ？

どちらのコントローラでもフォーマットは同じなので、エニュームは1つだけです。

```
enum DMA_MODE_REG_MASK {  
    DMA_MODE_MASK_SEL = 3,  
    DMA_MODE_MASK_TRA = 0xc,  
    DMA_MODE_SELF_TEST = 0,  
    DMA_MODE_READ_TRANSFER = 4,  
    DMA_MODE_WRITE_TRANSFER = 8,  
    DMA_MODE_MASK_AUTO = 0x10,  
    DMA_MODE_MASK_IDEC = 0x20,  
  
    DMA_MODE_MASK = 0xc0,  
    DMA_MODE_TRANSFER_ON_DEMAND = 0,
```

```

        DMA_MODE_TRANSFER_SINGLE = 0x40,
        DMA_MODE_TRANSFER_BLOCK = 0x80,
        DMA_MODE_TRANSFER CASCADE = 0xC0
    };
}

```

DMA0_MODE_REGが0x0b (DMA 0モードレジスタ)、DMA1_MODE_REGが0xd6 (2番目のDMAモードレジスタ) だとすると、特定のチャンネルのDMAモードを設定するために必要なことは以下の通りです。

```

void dma_set_mode (uint8_t channel, uint8_t mode)

{ int dma = (channel < 4) ? 0 : 1;
int chan = (dma==0) ? channel : channel-4;

dma_mask_channel (channel);
outportb ( (channel < 4) ? (DMA0_MODE_REG) : DMA1_MODE_REG, chan | (mode) );
dma_unmask_all (dma );
}

///! prepares channel for read
void dma_set_read (uint8_t channel) {

    dma_set_mode (channel,
                  DMA_MODE_READ_TRANSFER | DMA_MODE_TRANSFER_SINGLE | DMA_MODE_MASK_AUTO);
}

///! prepares channel for write
void dma_set_write (uint8_t channel) {

    dma_set_mode (channel,
                  DMA_MODE_WRITE_TRANSFER | DMA_MODE_TRANSFER_SINGLE | DMA_MODE_MASK_AUTO);
}

```

このルーチンでは、任意のチャンネルのモードを設定することができます。いいでしょう？例えば、フロッピードライブの書き込みを準備したい場合は、dma_set_mode (2, 0x5A)を実行します。(フロッピーはプライマリDMACのチャンネル2を使用していることを覚えていましたか?)と0x56

= 01010110のバイナリです。上のリストと比較すると、Mode=01(Single Transfer)、AutoInitが設定されている(完了後に自動初期化する)、転送タイプ=01(Write)、チャンネル2(10)となっています。

DMA_MODE_MASK_AUTOビットは便利なビットです。これにより、コントローラのリセットとチャンネルバッファのアドレスとカウントの設定を行うことで、DMACを最初から初期化することができます。このビットが設定されていない場合は、読み出しや書き込みのたびにDMACを再初期化する必要があります。

注：AutoInitビット(DMA_MODE_MASK_AUTO)は、Virtual PCではあまりサポートされていないようです。このため、Virtual PCでの移植性を維持するために、AUTOINITではなく、読み書きのたびにDMACを再初期化する方法を選択しました。他のエミュレータやマシンではサポートされていないかもしれません。

リクエストレジスタ（ライト）

このレジスタは、ソフトウェアがDMACに直接送信することを可能にします。最初の2ビットはチャンネルの選択に使用されます。例えば、00=チャンネル0、01=チャンネル1、10=チャンネル2、11=チャンネル3となります。3番目のビットは、0の場合、チャネル要求ビットをリセットします。1の場合は、リクエストビットを設定します。

ビット0-1：チャンネルセレクト0

ビット2：0=チャンネル要求ビットをリセット、1=要求ビットをセット

Request RegisterはMemory-to-Memoryの操作に使用されます。コマンドレジスタを思い出すと、i86アーキテクチャではMemory-to-Memoryトランザクションを有効にすることはできませんし、すべきでもありません。このため、このレジスタは重要ではありません。

チャンネルマスクレジスタ（ライト）

このレジスタでは、1つのDMAチャネルをマスクすることができます。ビット0と1でチャンネルを設定します (00=チャンネル0、01=チャンネル1、10=チャンネル2、11=チャンネル3)。ビット4は、チャネルをマスクするかアンマスクするかを決定します。ビット4が0であれば、チャネルのマスクを解除します。1であれば、マスクされます。他のビットはすべて未使用です。

ビット0-1：チャンネルセレクト

ビット2：0=unmasksチャンネル、1=masksチャンネル その

マスクレジスタ（ライト）

このレジスタには、どのチャンネルが現在マスクされているか、またマスクされていないかの情報が含まれています。この8ビットのレジスタの上位4ビットは常に未使用です。下位4ビットは、4つのチャンネルのいずれかをマスクまたはアンマスクするために使用されます。例えば、ビット0はチャンネル0、ビット1はチャンネル1、というようにです。注：カスケード接続により、チャンネル4をマスクすると、チャンネル4,5,6,7もマスクされます。

ビット0：チャンネルセレクト0
 ビット1：チャンネルセレクト1
 ビット2：チャンネルセレクト2

ビット3：チャンネルセレクト3
 その他のビットは未使用です。

例えば、任意のチャンネルをマスク（無効化）するルーチンを用意し、必要なのはそれぞれのビットを設定することだけです。

```
void dma_mask_channel(uint8_t channel)
{
    if (channel <= 4)
        outportb(DMA0_CHANMASK_REG, (1 << (channel-1)));
    else
        outportb(DMA1_CHANMASK_REG, (1 << (channel-5)));
}
```

同様に、チャネルのマスクを解除するには、そのビットをクリアすればよい。

```
void dma_unmask_channel (uint8_t channel)
{
    if (channel <= 4)
        outportb(DMA0_CHANMASK_REG, channel);
    else
        outportb(DMA1_CHANMASK_REG, channel);
}
```

これらのルーチンはいずれも、DMA0_CHANMASK_REGが0x0a（DMACマスクレジスタのi/oポート）、DMA1_CHANMASK_REGが0xD4（2つ目のDMACマスクレジスタのi/oポート）であることを前提としています。

複数のチャンネルを同時に設定できるため、このレジスターでは、複数のチャンネルを同時にマスクしたり、マスクを解除したりすることができます。

ステータスレジスタ

ステータスレジスタのフォーマットは以下の通りです。

ビット0：TC0 チャンネル0が転送完了（TC）に達した場合に設定 ビット1：TC1 チャンネル1が転送完了（TC）に達した場合に設定 ビット2：TC2 チャンネル2が転送完了（TC）に達した場合に設定 ビット3：TC3 チャンネル3が転送完了（TC）に達した場合に設定 ビット4：REQ0 チャンネル0がDMAリクエスト（DRQ）待ちの場合に設定 ビット5：REQ1 チャンネル1がDMAリクエスト（DRQ）待ちの場合に設定 チャンネル0がDMAリクエスト（DRQ）を保留している場合はREQ0を設定 ビット5：チャンネル1がDMAリクエスト（DRQ）を保留している場合はREQ1を設定 ビット6：チャンネル2がDMAリクエスト（DRQ）を保留している場合はREQ2を設定 ビット7：チャンネル3がDMAリクエスト（DRQ）を保留している場合はREQ3を設定

このレジスターはあまり有用ではありません。ほとんどの場合、DMACを制御しているデバイスは、転送が完了したときにIRQを送信するので、このレジスターをポーリングして情報を得る必要はありません。最初の4ビットは、そのチャンネルの転送が完了しているかどうかを示し、最後の4ビットは、そのチャンネルに保留中のDMAリクエストがあるかどうかを示します。

ISA DMAコマンド

コントローラには、ソフトウェアがコントローラにコマンドを送信できるようにするための特別なレジスターが用意されています。これらのコマンドは特定のビットフォーマットを必要とせず、簡単な入出力操作で起動することができます。

DMACは、アドレスバス（A0-A3ライン）のデータと、ORQおよびIOWラインのステータスによって、コマンドを認識します。これらのレジスターには特別なものは何もないことに注意してください。これらは、本章の冒頭にある汎用レジスターの表にも記載されています。

クリア バイト ポインタ フリップフロップ

これは特別なi/oアドレスポートで、8ビットDMAC（プライマリDMAC）で作業する際に、16ビット転送の間にフリップフロップを制御することができる。

両方のDMACに2つのポートがあります。

ISA DMAC Flip-Flop Ports	
Port	Description
0x0C	DMAC 0 (16 bit) Slave (write)
0xD8	DMAC 1 (8 bit) Master (write)

たとえば、DMA0_CLEARBYTE_FLIPFLOP_REGを0x0c、DMA1_CLEARBYTE_FLIPFLOP_REGを0xD8とすると、以下のルーチンでフリップフロップの設定またはクリアが行われます。

```
void dma_reset_flipflop(int
    dma){ if (dma < 2)
        return;

    //! it doesnt matter what is written to this register
    outportb( (dma==0) ? DMA0_CLEARBYTE_FLIPFLOP_REG : DMA1_CLEARBYTE_FLIPFLOP_REG, 0xff);
}
```

Reset

よく似た方法で、以下のレジスタに任意の値を書き込むことで、DMACをリセットすることができます。

ISA DMAC Reset Ports	
Port	Description
0x0D	DMAC 0 (16 bit) Slave (write)
0xD8	DMAC 1 (8-bit) Master (write)

例えば、DMA0_TEMP_REGを0x0Dと仮定した場合。

```
void dma_reset (int dma){

    //! it doesnt matter what is written to this register
    outportb(DMA0_TEMP_REG, 0xff);
}
```

Unmask All Registers

このコマンドも同様に、同じコンセプトで作られています。すべてのハードウェアプログラミングのコマンドがこのように簡単にできたらいいと思いませんか？）

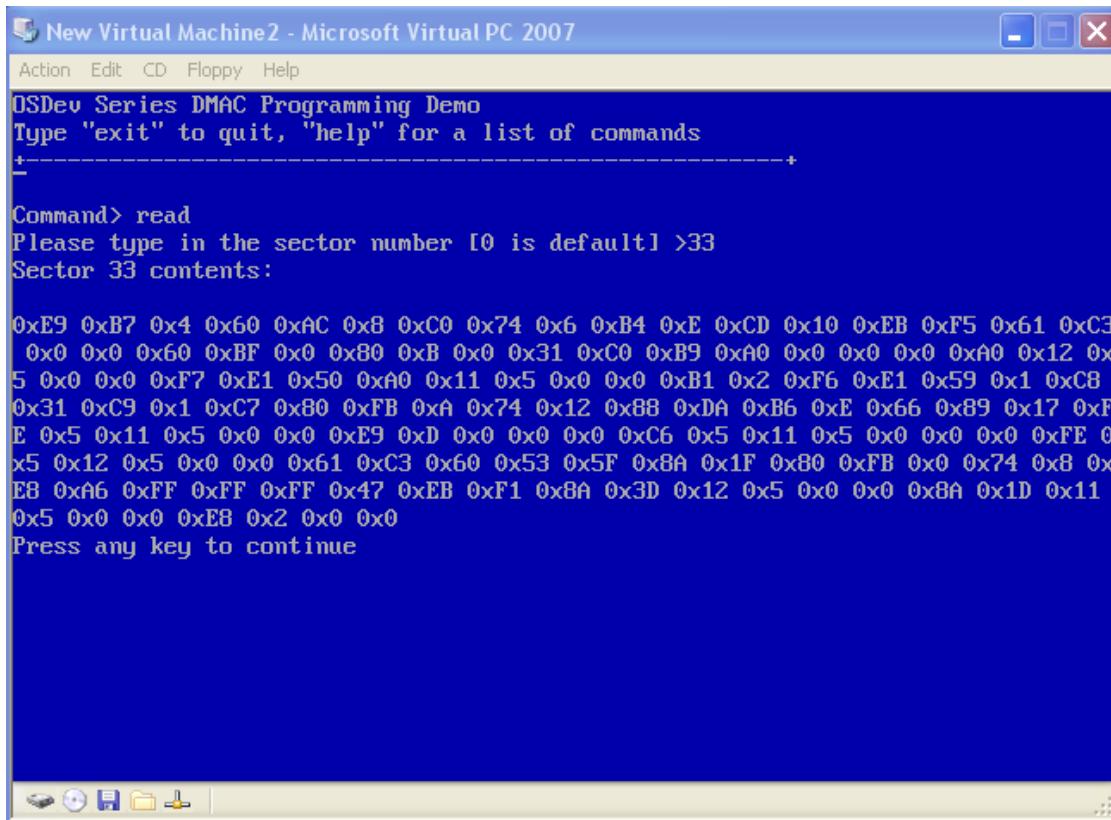
ISA DMAC UnMask All Ports	
Port	Description
0x0E	DMAC 0 (16 bit) Slave (write)
0xDC	DMAC 1 (8-bit) Master (write)

つまり、DMA1_UNMASK_ALL_REGが0x0Eだとすると、これはスレーブDMACからすべてのレジスタのマスクを解除することになります。

```
void dma_unmask_all (int dma){

    //! it doesnt matter what is written to this register
    outportb(DMA1_UNMASK_ALL_REG, 0xff);
}
```

Demo



*Virtual PCで動作するデモ D
emo ダウンロード*

やったー、またデモの時間だー 悪いニュースは、このデモは前の章と全く同じに見えるということです（ただし、BochsとVirtual PCの両方で動作するようになりました）。良いニュースは、このデモが新しいDMAインターフェースを使用するようにアップグレードされていることです。

新しいコードの核となる部分は、本章のすべてのコードを含むHAL - dma.hおよびdma.cppにあります。しかし、1つだけマイナーな変更点があります。DMACのModeレジスタのAUTOINITビットはVirtual PCではあまりサポートされていないので、dma_set_readおよびdma_set_writeルーチンは、デモコードではビットをセットしません。

```
///! prepares channel for read
void dma_set_read (uint8_t channel) {
    dma_set_mode (channel, DMA_MODE_READ_TRANSFER | DMA_MODE_TRANSFER_SINGLE);
}

///! prepares channel for write
void dma_set_write (uint8_t channel) {
    dma_set_mode (channel,
                  DMA_MODE_WRITE_TRANSFER | DMA_MODE_TRANSFER_SINGLE);
}
```

リードセクター操作の間、フロッピードライバーのfpydsk_read_sector_imp ルーチンは、DMAC を初期化し、DMAC をリード操作のために準備します。ルーチンの残りの部分（編集されています）は前章と同じで、FDCにREADコマンドを送信する役割を担っています。DMA_BUFFERは、DMACの転送に使用できるフリーメモリのバッファに過ぎません。dma_initialize_floppy は、新しい DMA ミニドライバーを使って DMAC を初期化し、フロッピードライバーが使用できるように準備します。（DMACを初期化した後、ドライバーのdma_set_readルーチンをチャンネルFDC_DMA_CHANNELに呼び出して、DMACのREAD操作の準備をします。FDC_DMA_CHANNELはチャンネル2である（FDCがDMACのチャンネル2を使用することを覚えているだろうか）。

```
///! read a sector
void flpydsk_read_sector_imp (uint8_t head, uint8_t track, uint8_t sector)

{ uint32_t st0, cyl;

  ///! initialize DMA
  dma_initialize_floppy ((uint8_t*) DMA_BUFFER, 512 );

  ///! set the DMA for read transfer
  dma_set_read ( FDC_DMA_CHANNEL );

  ///! rest of the code is the same...
}
```

`dma_initialize_floppy` は、新しいミニドライバーを使った DMAC をフロッピードライバーで使えるように準備する役割を果たします。ここからが楽しいことの始まりです。

まず、`dma_reset()`を呼び出してマスターDMACをリセットします。次に、チャンネル2 (FDCが使用する) をディスエーブル (マスク) するためには

`dma_mask_channel()`を使用します。これにより、チャンネルが使用されていないことを確認し、変更できるようになります。

さて、ここからが本番です。チャンネルが使用するアドレスを設定するには、`dma_set_address`ルーチンを呼び出します。これにより、チャンネルにアドレスの下位と上位を設定することができます。ユニオンを使用することで、メンバーのバイトコンポーネントへのアクセスが少し簡単になります。つまり、チャネルが使用するバッファに `a.l` を設定します。ユニオンのおかげで、`a.byte[0]`は値の下位バイトを、`byte[1]`は2番目のバイトを参照しています。バッファの大きさを表す`length`も同様にします。つまり、`dma_set_address`を呼び出してバッファのアドレスの下位バイトと上位バイトを設定し、同じように`dma_set_count`を呼び出して長さを設定します。いいでしょう？

わかった、わかった、それはそれでいいんだけど、`dma_reset_flipflop`の呼び出しはどうするの？フリップフロップは、8ビットDMACで16ビットのデータを扱うときのみ使用されます。もし16ビットDMACで作業していたら、このフリップフロップを呼び出す必要はありません。フリップフロップは、16ビットデータのハイバイトとローバイトを選択するのに使われます。フリップフロップをリセットすると、次のバイトデータがローバイトになることをDMACに伝えることになります。フリップフロップがデフォルトの位置にない場合は、ハイバイトとして選択されます。DMACは8ビットのデータバスで16ビットのデータを扱うので、これを選択する必要があります。このバイトが16ビットデータのどの部分を指しているのか、どうやって知ることができるのでしょうか？

最後に、`dma_set_read()`を呼び出してDMACを読み取り操作用に設定し、すべてのチャンネルのマスクを解除して、再びデバイスが使用できるようにする。これは、FDCがDMACのチャンネル2を使用できるようにするために重要である。

```
bool __cdecl dma_initialize_floppy(uint8_t* buffer, unsigned length){ union{ uint8_t byte[4];//Lo[0], Mid[1], Hi[2] unsigned long l; }a, c;

a.l=(unsigned)buffer;
c.l=(unsigned)length-1;

//Check for buffer issues
if ((a.l >> 24) || (c.l >> 16) || (((a.l & 0xffff)+c.l) >> 16)){ #ifdef _DEBUG
    _asm{
        mov      eax, 0x1337
        cli
        hlt
    }
#endif
    return false;
}

dma_reset (1);
dma_mask_channel( FDC_DMA_CHANNEL );//Mask channel 2
dma_reset_flipflop ( 1 );//Flipflop reset on DMA 1

dma_set_address( FDC_DMA_CHANNEL, a.byte[0],a.byte[1]);//Buffer address
dma_reset_flipflop( 1 );//Flipflop reset on DMA 1

dma_set_count( FDC_DMA_CHANNEL, c.byte[0],c.byte[1]);//Set count
dma_set_read ( FDC_DMA_CHANNEL );

dma_unmask_all( 1 );//Unmask channel 2

return true;
}
```

結論

さて、もう1つの章が終わりましたね。この章は、これまでの章ほど複雑で難しいものではなかったので、いい息抜きになったのではないでしょうか？

ここから先に進むには、ディスクからファイルを読み込む機能が必要です。データをディスクから読み込む機能はありますが、ファイルはありません。これにはファイルシステムドライバが必要です。しかし、待ってください！FAT12については、すでに2回ほど取り上げています。書き直しがいかに多いかということですね。今回は、同じ内容を3回も繰り返すのではなく、もう1つのテーマを追加します。仮想ファ

イルシステム(VFS)です。次の章では、ちょっとしたお遊びとして、デモプログラムを実行する機能も追加するかもしれません。)

オペレーティングシステムにグラフィカルなタッチを加えたいという読者が多いことから、VBE (Vesa Bios Extensions) やVGA (Video Graphics Array) ／SVGA (Super VGA) に関連するいくつかの上級編も公開する予定です。また、現在のシステムを本物のマイクロカーネルにするために、DLLのサポート、ドライバ、ネイティブPEリソースのサポートなども行います。

これからも素敵なことがたくさん待っています。) もし、あなたが私たちに取り上げてほしいトピックのアイデアがあれば、遠慮なく私に教えてください。

次の機会まで。
~マイク

BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中 質問やコメントは？お気軽にお連絡ください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。

リファレンス

82C37A CMOS High Performance Programmable DMA Controller データシート

"The Undocumented PC"

ご質問やご意見がございましたら お気軽にお問い合わせください。



Chapter 20

Home

Chapter 22

第



20章 ホーム 第22章



オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - ファイルシステムとVFS

by Mike, 2010

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

はじめに

OS開発のための終わりなきシリーズ、第22章へようこそ これは第22章というよりも、OS開発シリーズの2年目にあたります。

今回もまたファイルシステム関連のチュートリアルです（心配しないでください、これが最後です；）。最初のものは、ブートコードからメインのブートローダプログラムをロードするために必要でした。2番目のものは、メインのブートプログラムがカーネルをロードするために必要でした。今度は、カーネルがプログラムをロードして実行できるように、カーネル用にもう一つ必要です。しかし、この章と他の2つの章には違いがあります。この章はアセンブリ言語ではなくC言語で行われます。）

しかし、今回は新たな試みとして、仮想ファイルシステム（VFS）にも注目します。これにより、あらゆるファイルシステムドライバーと異なるディスクデバイスを同じように扱うことができるようになります。これは、ローカルディスクドライブだけでなく、あらゆるネットワークファイルシステムとのインターフェースにも使用できます。

いいですか？

ファイルシステム

アブストラクト

ファイルシステム

ファイルシステムは、情報を読み書きするための論理的な方法を定義するものです。ファイルシステムは、情報を読み書きするための論理的な方法を定義したもので、仕様の一つと言えます。ほとんどのPCのファイルシステムは、デスクトップの「ファイル」と「フォルダ」の概念に基づいています。

ファイルシステムには様々な種類があります。広く使われているもの（FAT12、FAT16、FAT32、NTFS、ext（Linux）、HFS（古いMACで使われている）など）もあれば、特定の企業が社内でのみ使用するファイルシステム（GFS - Google File Systemなど）もあります。他のファイルシステムは、特定の企業が社内でのみ使用するものです。また、独自のファイルシステムを開発・設計することもできます。

ファイルシステムは、データの保存と整理のために使用されます。ファイルシステムは、リムーバブルメディア（フロッピー、フラッシュドライブ、CD、DVD）、ローカルドライブ（ハードディスクドライブ）、ネットワーククライアント上のファイルやディレクトリに簡単にアクセスする方法を提供します。ファイルシステムは、インメモリーイメージとしても存在します。例えば、特別なタイプのファイルシステムの「足型」を含んだファイルを読み込むことができます。

ファイルとフォルダー

ファイルとは、プログラムやユーザーにとって何かを表すデータの集まりです。このデータは、私たちが望むものであれば何でもあります。それは、データをどのように解釈するかによります。例えば、テキストファイルは、テキスト情報を含んでいます。ファイルは、何かのイメージであることもあります。フォルダは、ファイルを論理的にまとめたものです。ディレクトリとも呼ばれています。

ディレクトリは、大量のファイルを管理する手段です。ディレクトリは通常、ツリー構造になっています。これをディレクトリツリーと呼びます。すべてのディレクトリとファイルの親となるディレクトリは1つだけで、ルートディレクトリです。ファイルパスは、ディレクトリツリー内のファイルの位置を示します。例えば、「a:myfile.txt」というファイルは、myfile.txtがファイル名です。a:myfile.txtは「mydir」というサブディレクトリーにある「myfile.txt」というファイルで、「a:」というデバイスのルートディレクトリーにあります。

ファイルとフォルダーのネーミング

フォルダやファイルの名前は、そのファイルやフォルダを表す文字列であり、通常はその内容によって表される。ファイルシステムでは、ファイル名とフォルダ名の意味合いが異なり、それぞれに制約があります。例えば、FAT12では、ディレクトリエントリのファイル名やフォルダ名を11バイトの配列（ファイル名が8、拡張子が3。これは「8.3命名規則」とも呼ばれています。）これにより、ファイル名とフォルダ名は11文字に制限されます。一方、NTFSはLFN（Long File Name）をサポートしており、255文字に制限されています。また、NTFSでは、ファイル名をファイル属性と

ともにマスターファイルテーブルに格納しています。

ほとんどのファイルシステムのファイル名は、大文字と小文字を区別しません。しかし、ファイルシステムによっては、ファイル名を内部的に異なる方法で保存している場合があります。例えば、フロッピーディスク上のファイルに8.3の小文字のファイル名を付けていても、OSからそのファイルを読み込むときにはすべて大文字のファイル名を使うことができるご存知でしょうか。Windowsはファイル名のLFNを表示しますが、FAT12の8.3ファイルのエントリには、その8.3の全大文字のファイル名だけが表示されます。これが可能になるのです。

ファイルの種類

シンボリックリンクの

シンボリックリンクとは、パスを短くするための方法です。例えば、a:/folder/link.lnkは、a:/otherfolder/subfolder/yet another folder/link.txtを指します。これで、テキストファイルに簡単にアクセスできます。また、シンボリックリンクは

フォルダを整理するためによく使われます。Windowsのスタートメニューのようなものです。プログラムへのシンボリックリンクが含まれています。シンボリックリンクを実装するのはそれほど難しくありません。指定されたノード（リンク）を見つけます。それはリンクのように見えるので、実際のパスを取得して、代わりにそのファイルを読みます。

Windowsショートカットは、シンボリックリンクの一種です。

パイプ

IPC (InterProcess Communication) の一種である「パイプ」。パイプとは、通常、2つ以上のプロセス間に存在する仮想ファイルのことです。最も良い例は、Unixのstdout、stdin、stderrでしょう。これらは通常のファイルとして扱われますが、stdoutに書き込まれたデータは画面に表示されます（またはstdout.txtに表示されます）。

特殊なファイルタイプ

メタファイル

ファイルシステムの中には、ファイルシステム専用の特別なファイルやフォルダを実装しているものもあります。通常、同じディレクトリに同じ名前のファイルやフォルダを2つ置くことはできません（フォルダと同じ名前のファイル名も置けません）。そのため、これらの隠しファイルを使ってファイルやフォルダの名前を付けることも、実装によってはできない場合があります。

例えば、NTFSではファイルシステム用にいくつかのメタファイルが用意されています。これらのファイルは、システムドライブのルートディレクトリ（通常はC:）に配置されます。例えば、\$MFT、\$MFTMirr、\$LogFileなどのファイルがあります。隠しファイルの表示やシステムファイルの表示にチェックを入れても表示されることはありませんが、ここに上記のような名前のファイルを作成するとどうなるかを見てみましょう。これらのファイルは他の場所でも作成できますが、ルートディレクトリに作成すると、メタファイルのために「ファイルが既に存在します」というエラーが発生します。

デバイスファイル

Unix系システム、DOS（ひいてはWindows）には、デバイスを表す特殊な「ファイル」であるデバイスファイルがあります。例えば、NUL(ヌルデバイス)、CLOCK\$、PRN(プリンタ)などです。ここでは、デバイスファイルの一覧を紹介します。

```
con prn
aux
clock$ n
ul

com0, com1, ... com9 lpt0,
lpt1, ... LPT9
```

これらの名前はDOSやWindowsでは特別な意味を持つため、ファイルやフォルダに上記のような名前をつけることはできません。

…としています。

は、一部のファイルシステムで採用されている特殊なファイルです。は、カレントディレクトリを参照するファイル情報を含むファイルのファイル名である。は、そのファイルの親ディレクトリを参照する情報を含むファイルのファイル名で、「...」は、そのファイルの親ディレクトリを参照する情報を含むファイルのファイル名である。例えば、c:\mydir\file.txtというファイルがあり、c:\mydirがカレントディレクトリだった場合、パス名...はC:を、パス名...はc:を参照します。

ファイルシステムの種類

フラットファイルシステム

フラットファイルシステムとは、サブディレクトリを持たないファイルシステムである。その代わり、すべてのファイルは同じ（ルート）ディレクトリにある。初期のコンピュータシステムの多くはフラットファイルシステムを採用していた。最近のOSでは、より高度な階層型ファイルシステムが採用されている。フラットファイルシステムは、小型で導入しやすい反面、新規に開発するのが難しい。

階層型ファイルシステム

このタイプのファイルシステムは、サブディレクトリをサポートしています。最近のほとんどのファイルシステム(FAT12,FAT16,FAT32,etx,NTFSを含む)はこのカテゴリーに当てはまる。(FAT12の最初のバージョンは、フラット・ファイル・システムでした。しかし、その後のバージョンではサブディレクトリをサポートしています)。

ジャーナリング・ファイルシステム

このタイプのファイルシステムでは、ファイルシステムの変更の「ジャーナル」を使用します。これは、システムがファイルやディレクトリに加えようとする変更を、手順を完了する前に記録したものです。これにより、ファイルシステムの操作（ファイルの書き込みなど）中にクラッシュが発生した場合、ジャーナルを読んで変更を元に戻し、ファイルシステムを修復することができます。

ファイルシステムドライバ

ファイルシステムが「ファイル」や「ディレクトリ」を読み書きするための仕様を定めたものであるのに対し、ファイルシステムドライバーは、特定の種類のファイルシステムを実装するためのものです。ファイルシステムドライバーの例としては、マイクロソフト社のNTFSファイルシステムを実装したntfs.sysがあります。また、ファイルシステムドライバは、大規模なソフトウェアの中にミニドライバとして実装されることもあります。ブートローダがその良い例です。ブートローダは、別のドライバプログラムなしでディスクからファイルをロードできる必要があるため、ブートローダ自体の中に、さまざまな種類のファイルシステムに対応した複数のファイルシステムミニドライバが含まれています。シリーズでブートローダを開発された方は、すでにFAT12ファイルシステムを経験され、当社のブートローダ用にFAT12ミニドライバを開発されています。

仮想ファイルシステム (VFS)

アブストラクト

仮想ファイルシステム(VFS)は、特定のファイルシステム実装の上にある抽象化レイヤーです。ソフトウェアは、VFSを介してストレージデバイスにアクセスします。これにより、ソフトウェアは、使用されているデバイスやファイルシステムの知識がなくても、異なるストレージデバイスに読み書きすることができます。また、インストールされたファイルシステムやデバイスの数に関わらず、同じコードで動作させることができます。基本的な考え方は、単一のシステムインターフェースであらゆるファイルシステムを統一的に扱えるようにすることです。Windows、Linux、Mac OSはそれぞれ異なる方法でVFSをサポートしています。

インプリメンテーション

VFSを導入するにはさまざまな方法があります。

マウントポイント一覧

マウントポイントリストとは、マウントされているファイルシステムとそのマウントされている場所のリストです。例えば、ファイルを読み取る必要がある場合、OSは通常、VFSのReadFile()関数を呼び出し、マウントされたファイルシステムのリストを検索して、ファイルが入っているデバイスとファイルシステムを見つけます。そして、そのファイルシステムのReadFile()関数に読み込み要求を渡します。

ノードグラフ

ノードグラフは、ファイル、フォルダ、マウントポイントなど、さまざまなタイプのファイルを表すノードのグラフを含んでいます。各ファイルノード構造には、通常、ファイルを読み書きするためのファイルシステム固有のルーチンへの関数ポインタが含まれています。

例えば、次のようなFILE構造を作ることができます。

```
typedef struct _FILE {
    char          name[32];      //filename
    uint32_t      flags;         //flags
    uint32_t      fileLength;   //length of file
    read_func     read;          //function pointers to read, write, open, close file
    write_func    write;
    open_func     open;
    close_func   close;
}FILE, *PFILE;
```

関数のポインタがこのFILE構造体に格納されていることに注目してください。例えば、ファイルを読みたいので、fopen()を呼び出し、最終的にVFSのOpenFile()関数を呼び出します。VFSのファイル操作ルーチンが行う必要があるのは、特定のFILEの関数ポインタに制御を渡すことだけです。

```
void VfsOpenFile (PFILE file, const char* filename) { if
    (file)
        file->open (filename);
}
```

これにより、ファイルシステムで定義されたルーチンを呼び出すことができます。

DOSとWindows

DOSおよびWindowsでは、マウントされたファイルシステムを表すために、「a」から「z」までの文字が割り当てられています。Windowsでは、ドライブレターとそのオブジェクトマネージャ名の間にシンボリックリンクが張られます。例えば、ドライブレター「c」の場合。例えば、ドライブレターc:(symbolic link name ¥¥C:)は、オブジェクト名¥Device¥HardDiskVolume1のデバイスオブジェクトにマッピングされます。ファイルシステムは、デバイスオブジェクトを所有するために自らを登録することができる。オブジェクトを所有しているファイルシステムが見つかった場合、ファイルパス名の残りの部分（この例では "myfile.txt"）は、そのファイルシステムのFileOpen()関数に渡されます。

ドライブレターの割り当て

Windowsは、マウントされたファイルシステムを表すデバイスやパーティションへのドライブレターの割り当てをサポートしています。（起動時に、デバイスオブジェクトを所有するファイルシステムドライバーが登録されていない場合、Windowsはデバイスに対してRAWミニドライバーを使用します）。ドライブレターは、ネットワーク共有ドライブ、仮想ディスクイメージ、ローカルまたはネットワーククライアント内の別の場所へのシンボリックリンクを参照することもできます。また、ドライブレターは、ネットワーク共有ドライブ、仮想ディスクイメージ、ローカルまたはネットワーククライアント内の別の場所へのシンボリックリンクを指すこともあります。

インターフェース

ここでは、VFSの実装を簡単にするために、ドライブレターの割り当てとマウントポイントのリストを使用します。このシリーズで紹介するOSには、デバイス管理やI/O管理がないので、シンプルな実装にする必要があります。

個人的には、ファイルシステムドライバよりもVFSを先に開発することをお勧めします。そうすれば、VFSのインターフェースやフレームワークがすでに完成しているからです。

FILE

C言語を使ったことがある人なら、悪名高いFILE*データ型を知っているでしょう。FILE*は、ファイルオブジェクトへのポインタを表す抽象データ型(ADT)です。ISO Cでは、C言語の実装において、FILE型を定義しなければならないと規定されていますが、どのようなものかは定義されていません。

は、構造体の内部にある。つまり、FILE*がISO Cであるのに対し、構造体の中身はインプリメンテーションで定義されている。ファイルの現在の状態を表すファイル構造を、好きなように定義することができます。ファイルには名前とサイズがあり、これはすでに2つのメンバーです。さらに、ファイルの終端 (EOF) を示すフラグや、ファイル固有のフラグが必要で、そのために2つのメンバーが必要です。また、ファイルの現在の位置（クラスターとクラスターのオフセット）を追跡する方法も必要で、現在は次のようにになっています。

```
typedef struct _FILE {
    char        name[32];
    uint32_t    flags;
    uint32_t    fileLength;
    uint32_t    id;
    uint32_t    eof;
    uint32_t    position;
    uint32_t    currentCluster;
    uint32_t    device;
}FILE, *PFILE;
```

簡単だったでしょう？ idは、必要に応じて識別のために使用することができます。

ファイルの種類

これまで説明してきたファイルには、ファイル、ディレクトリ、シンボリックリンクなど、さまざまな種類があります。ここでは、わかりやすくするために、ファイルとディレクトリにのみ注目します。これらは、ファイルの種類を表すために、上記のFILE構造のflagsメンバーで使用されます。

```
#define FS_FILE      0
#define FS_DIRECTORY  1
#define FS_INVALID    2
```

オペレーション

ファイルに対して行うことのできる代表的な操作があります。

開く

Close

Read

Write

Mount

Unmount

ファイルオブジェクト（ファイルやディレクトリなど、ファイルの種類は問わない）のオープンとクローズを行うのがOpenとClose、ファイルの種類に応じた読み込みと書き込みを行うのがReadとWriteである。これらはすべて、標準的なCファイルI/O関数を通じてプログラマに公開されている。

VFSでは、fsys.hにあるVolume Managerを通じて公開されています。

```
extern FILE volOpenFile (const char* fname);
extern void volReadFile (PFILE file, unsigned char* Buffer, unsigned int Length); extern
void volCloseFile (PFILE file);
extern void volRegisterFileSystem (PFILESYSTEM, unsigned int deviceID);
extern void volUnregisterFileSystem (PFILESYSTEM);
extern void volUnregisterFileSystemByID (unsigned int deviceID);
```

例えば、C言語のfopen()ルーチンを呼び出したとします。このルーチンは、FILEオブジェクトを返すvolOpenFile()ルーチンを呼び出します。ここでは「a:myfile.txt」のようにファイルへのパスを渡しました。ボリュームマネージャは、マウントポイントリストを調べ、「a」を表すデバイスIDにファイルシステムが登録されているかどうかを確認します。ファイルシステムが登録されていれば、そのファイルシステムドライバのFileOpen()メソッドに「myfile.txt」を渡して呼び出します。複雑に聞こえるかもしれません、ご心配なく。しかし、デモでの実装方法は非常に簡単です。

ボリュームマネージャーの導入

ファイルシステムの抽象化

まず必要なのは、ファイルシステム固有の情報を抽象化する方法です。これには、ファイルシステムの名前や、ファイルに対して実行可能な操作が含まれます。これには、関数ポインタを使用します。

```
char Name [8];
FILE (*Directory) (const char* DirectoryName);
```

```
void          (*Mount)()
void          (*Read) (PFILE file, unsigned char* Buffer, unsigned int Length);
void          (*Close) (PFILE);
FILE          (*Open) (const char* FileName);
```

インプリメンテーション

ボリュームマネージャは、デモのVFSを実装しています。ファイル fsys.h と fsys.cpp の中にあります。ドライブレターの割り当てを使ってデバイスを表現することを覚えていますか？26個のデバイスが存在するので、DEVICE_MAXという定数を作成しておくと便利です。各デバイスは1つのマウント可能なファイルシステムしか持てないので、それらをリスト（マウントポイントリストのようなもの）にして保存します。

```
#define DEVICE_MAX 26
//! File system list
PFILESYSTEM _FileSystems[DEVICE_MAX];
```

その仕組みは以下の通りです。ファイルシステムをポインタのリストとして保存しているので、ポインタが有効であれば、ファイルシステムがそこに登録されていることになります。配列の各要素は、それが参照するドライブレターを表しています。つまり、「a」は _FileSystems[0] に、「b」は

_FileSystems[1] などがあります。ファイルシステムは、自分が書き込んでいるディスクを管理する責任があります。

このメソッドを使用すると、非常に基本的ですが、簡単にデバイスにアクセスすることができます。例えば、volOpenFile() は、パスの最初の文字（ドライブレター）をチェックして、そのデバイスにファイルシステムが登録されているかどうか、リストを検索するだけです。登録されていれば、そのファイルシステムのopen() メソッドを呼び出して、ファイル名をドライブに渡すことができます。デフォルトでは「a」を使用しますが、入力パスに「:」が含まれている場合は、代わりにデバイスの最初の文字を使用します。これにより、volOpenFile を2つの方法で呼び出すことができます。“myfile.txt”のような文字列を渡す場合と、“a:myfile.txt”を渡す場合です。“a”はファイルが入っているデバイスです。かっこいいでしょ？

```
FILE volOpenFile (const char* fname) { if
    (fname) {
        //! default to device 'a'
        unsigned char device = 'a';
        //! filename
        char* filename = (char*) fname;
        //! in all cases, if fname[1]==':' then the first character must be device letter
        if (fname[1]==':') {
            device = fname[0];
            filename += 2; //strip it from pathname
        }
        //! call filesystem
        if (_FileSystems [device - 'a']) {
            //! set volume specific information and return file
            FILE file = _FileSystems[device - 'a']->Open (filename);
            file.deviceID = device;
            return file;
        }
    }
    FILE file;
    file.flags = FS_INVALID;
    return file;
}
```

その他のファイル操作ルーチンは基本的にすべて同じです。VFSがどのようにファイルシステムを保存しているかを知ると、volRegisterFileSystem() ファミリーのルーチンがどのように動作するかを推測できるでしょう。これらのルーチンが基本的に実行することは、ファイルシステムへのポインタをリストに格納することです。

```
void volRegisterFileSystem (PFILESYSTEM fsys, unsigned int deviceID)
{ if (deviceID < DEVICE_MAX)
    if (fsys)
        _FileSystems[deviceID] = fsys;
}
```

さて、それでは。そこで、ファイルシステムドライブを初期化し、VolRegisterFileSystem() を呼び出して自分自身を登録します。fopen() を呼び出し、それが VolOpenFile() を呼び出し、さらにそれがファイルシステムの open() メソッドを呼び出します。これですべてが整いましたが、何かが足りません。非常に重要なことがあります。

そうですね、私たちはそれに入るべきでしょう……もう一度……。

FAT12 - テイクスリー

はじめに

シリーズを通して、過去に2回、FAT12を見て、実装しました。そのため、今回も FAT12 を詳細に取り上げるつもりはありません。しかし、今

回はFAT12の復習を兼ねて、Cドライバのコードとその動作を紹介します。

必要に応じて第11章を参照しながらお読みください。

ブートセクター

重要なファイルシステム情報の多くが、ブートストラッププログラムとともにブートセクターに保存されていることを覚えていますか？具体的には、ブートセクタ内のPBP（Bios Paramater Block）に格納されています。

ファイルシステムをマウントする際には、BPBから情報を読み取り、後で使用するためにこの情報を保存する必要があります。そのためには、ブートセクタに一致する構造を作ればいいのです。

```
uint8_t           Ignore[3];          //first 3 bytes are ignored (our jmp instruction)
BIOSPARAMATERBLOCK Bpb;             //BPB structure
BIOSPARAMATERBLOCKEXT BpbExt;       //extended BPB info
uint8_t           Filler[448];        //needed to make struct 512 bytes
```

ブートセクターがどのように見えるかの良い例として、Stage1のブートローダプログラムがメモリ上でどのように見えるかを考えてみましょう。Stage1の一番最初の命令（第4章のデモ、Stage1.asmを参照）はjmp loaderでした。これは3バイトの命令なので、上の構造の最初の3バイトは、jmp命令のオペレーションコード（OPCode）です。

また、第4章では、OEM Paramater Block（別名：Bios Paramater Block（BPB））について説明しました。BPBは、3バイトのジャンプ命令の直後にあります。このため、BIOSPARAMATERBLOCKはこの構造の中で次の位置にあります。また、FAT32などの他のファイル・システム用にBPBを拡張したBIOSPARAMATERBLOCKEXT構造体も提供しています。

ブートセクタの最後の448バイトには、ブートセクタの残りのプログラムコードが含まれています。今のところ重要ではないので、フィラー・メンバーのパディングとして処理しています。これにより、BOOTSECTOR構造がディスク上のブート・セクター(512バイト)と正確に同じサイズになることが保証されます。

BIOSPARAMATERBLOCKは、BPBのフォーマットを定義する構造体です。第5章で詳しく説明しましたが、ブートセクターと同じ構造です。

```
typedef struct _BIOS_PARAMATER_BLOCK {
    uint8_t           OEMName[8];
    uint16_t          BytesPerSector;
    uint8_t           SectorsPerCluster;
    uint16_t          ReservedSectors;
    uint8_t           NumberOfFats;
    uint16_t          NumDirEntries;
    uint16_t          NumSectors;
    uint8_t           Media;
    uint16_t          SectorsPerFat;
    uint16_t          SectorsPerTrack;
    uint16_t          HeadsPerCyl;
    uint32_t          HiddenSectors;
    uint32_t          LongSectors;
} BIOSPARAMATERBLOCK, *PBIOSPARAMATERBLOCK;
```

上記の構造は、より親しみやすいものになっているはずです :) そうでない場合は、第5章の説明をお読みください。

しかし、BIOSPARAMATERBLOCKEXTは、新しいかもしれません。BPBについてはすでに詳しく説明し、過去にFAT12の解析に使用しましたが、FAT12のブートセクタはBPBの拡張メンバーに依存していません。しかし、FAT32はそうである。

```
uint32_t          SectorsPerFat32;      //sectors per FAT
uint16_t          Flags;                //flags
uint16_t          Version;              //version
uint32_t          RootCluster;         //starting root directory
uint16_t          InfoCluster;
uint16_t          BackupBoot;          //location of bootsector copy
uint16_t          Reserved[6];
```

これですべてです :) これらの構造体は、ファイルシステムドライバがBPB内のデータを参照するための簡単な方法です。これらの構造体は、ファイルシステムドライバがBPB内のデータを参照し、後にファイルシステムで使用するための簡単な方法を提供します。あとは、ブートセクタを読み込んで、PBOOTSECTORを介してデータにアクセスするだけです。)

前章で開発したフロッピーディスクのドライバーを使って、セクタを読み取る。

```
//! Boot sector info
PBOOTSECTOR bootsector;

//! read boot sector
bootsector = (PBOOTSECTOR) f1pydsk_read_sector (0);
```

必要なのはそれだけです :) 重要な情報はすべて bootsector.bpb に入っています。あとは、ファイルシステムをマウントするだけです...

ファイルシステムのマウント

BPBの情報がメモリに入ったところで、ファイルシステムを使用するための準備をする必要があります。まず、必要な情報を決めるところから始めます。

さて、ディスク上の総セクタ数を知る必要があります。また、ディレクトリエントリの総数も必要です。他にも、ファイルアロケーションテーブル (FAT) やルートディレクトリの使用に役立つ情報があります。

```
typedef struct _MOUNT_INFO {
    uint32_t numSectors;
    uint32_t fatOffset;
    uint32_t numRootEntries;
    uint32_t rootOffset;
    uint32_t rootSize;
    uint32_t fatSize; uint32_t
    fatEntrySize;
} MOUNT_INFO, *PMOUNT_INFO;
```

さて...。BOOTSECTOR構造体にブートセクタが格納されているのを覚えていませんか？これを利用して、BPBの情報の一部をMOUNT_INFO構造体にコピーすればよいのです。

さてさて。FAT12フォーマットのディスクで、最初のFATとルートディレクトリの位置を確認してみましょう。

Boot Sector	Extra Reserved Sectors	File Allocation Table 1	File Allocation Table 2	Root Directory (FAT12/FAT16 Only)	Data Region containing files and directories.
-------------	------------------------	-------------------------	-------------------------	-----------------------------------	---

2つのFATがあることに注目。最初のFATは、ディスクのブートセクタの直後になります。このため、MOUNT_INFOのfatOffsetを1に設定しています。また、Root Directoryは両方のFATの直後にあることに注意してください。これを知っていると、ルート・ディレクトリの開始セクタを求める簡単な計算ができます。(NumberofFATs * sectorsPerFAT) + 1。ブートセクタのために1を加える必要があります。

これで、最初のFATとルートディレクトリの位置がわかりました。ルート・ディレクトリのサイズを求めるために必要なのは、ルート・ディレクトリのエントリ数と各エントリのサイズです。FAT12の各ディレクトリエントリは、32バイトのサイズを持つ特定の構造形式です。したがって、必要なのはブートセクタ->Bpb.NumDirEntries * 32です。これは、ディレクトリが占めるバイト数です。これをセクタあたりのバイト数で割って、セクタ数に変換します。

```
//! store mount info
_MountInfo.numSectors      = bootsector->Bpb.NumSectors;
_MountInfo.fatOffset        = 1;
_MountInfo.fatSize          = bootsector->Bpb.SectorsPerFat;
_MountInfo.fatEntrySize     = 8;
_MountInfo.numRootEntries   = bootsector->Bpb.NumDirEntries;
_MountInfo.rootOffset       = (bootsector->Bpb.NumberOfFats * bootsector->Bpb.SectorsPerFat) + 1;
_MountInfo.rootSize         = (bootsector->Bpb.NumDirEntries * 32) / bootsector->Bpb.BytesPerSector;
```

それだけではありません。ここでFAT12ドライバーの初期化は完了です。簡単でしょう？重要なファイルシステム情報はMOUNT_INFOに入っているので、あとはディレクトリを解析してファイルをロードするだけです。)

ディレクトリの解析

フォーマット

FAT12のディレクトリは、ファイルやサブディレクトリの情報を提供する32バイトの構造体で構成されています。各ディレクトリ・エントリは以下の形式を持っています。

```
uint8_t  Filename[8];           //filename
uint8_t  Ext[3];               //extension (8.3 filename format)
uint8_t  Attrib;               //file attributes
uint8_t  Reserved;
uint8_t  TimeCreatedMs;        //creation time
uint16_t TimeCreated;          //creation date
uint16_t DateCreated;          //creation date
uint16_t DateLastAccessed;
uint16_t FirstClusterHiBytes;
uint16_t LastModTime;          //last modification date/time
uint16_t LastModDate;
uint16_t FirstCluster;          //first cluster of file data
uint32_t FileSize;             //size in bytes
```

それがすべてです :) これはディレクトリエントリです。私たちのDIRECTORY構造に格納されている情報は、サブディレクトリであったり、ファイルであったりします。FilenameとExtは、ファイルまたはディレクトリの8.3フォーマット名を含む。

Attribは、ファイルやディレクトリの属性を含みます。参考までに以下のような値があります。読み

Volume Labelです。8

サブディレクトリー

0x10 アーカイブ 0x20

デバイス: 0x60

これは必要ないので、シリーズでは使用しませんのでご了承ください。ただし、お好きな方はご自身のシステムでファイルの属性を操作・設定するサポートを行うことができます。

この構造体のすべての日付メンバーは、特定のビットフォーマットに従います。

ビット0~4: 曜日 (0~31)

ビット5-8: 月 (0-12)

ビット9-15: 年

この構造のすべてのタイムメンバーは、特定のビットフォーマットに従います。

ビット0~4: セカンド

ビット5-10分

ビット11-15: Hour

本連載では、ファイルやディレクトリの日時情報を変更したり取得したりする必要がないため、これらを使用していません。しかし、読者の皆様には、後からでもお好きなように機能を追加していただきたいと思います。

FAT12 フォーマットのフロッピーディスクでは、クラスタは1セクタ(512バイト)と同じ大きさであることを覚えておいてください。このため、DIRECTORYのFirstClusterフィールドは、ファイルの最初のセクタを指す。したがって、このセクタを読み取ることで、ファイルの最初の512バイトを効果的に読み取ることができます。

それでは、ディレクトリを解析してファイルを探してみましょう。

パーシング

ディレクトリには、ディレクトリエントリ構造のリストが含まれていることを覚えておいてください。このことを知っていれば、ディレクトリを解析してファイルやディレクトリを見つけることがとても簡単になります。

まず、ルートディレクトリの読み込みから始めます。ファイルシステムのマウント時にBPBからルートディレクトリのセクタを取得し、_MountInfo.rootOffsetに格納したこと思い出してください。したがって、必要なのはセクターをロードして、DIRECTORY*を使ってディレクトリエントリにアクセスすることだけです。

そして、ファイル名をループして比較し、一致するものを探します。ToDosFileName()を使って、入力ファイル名をDOS 8.3のファイル名形式に変換します。例えば、入力ファイル名「myfile.txt」をFAT12内部フォーマット「MYFILE TXT」に変える。

セクターを読み込んで、セクター内の各エントリを比較します。また、ファイル名が一致するかどうかを単純なstrcmp()呼び出しでテストできるよう、ファイル名をC言語の文字列に変換していることに気づくでしょう。一致するものが見つかったら、FILE構造体に記入してそれを返します。見てみましょう。

```
FILE fsysFatDirectory (const char* DirectoryName)
{
    FILE file;
    unsigned char* buf;
    PDIRECTORY directory;

    //! get 8.3 directory name
    char DosFileName[11];
    ToDosFileName (DirectoryName, DosFileName, 11);
    DosFileName[11]=0;
```

DirectoryNameには、検索したいディレクトリやファイル名が入ります。myfile.txtのような入力ファイル名を、DOS 8.3ファイルシステムのフォーマットである「MYFILE TXT」に変換し、DosFileNameに格納します。

```
for (int sector=0; sector<14; sector++) {
    //! read in sector
    buf = (unsigned char*) flpydsk_read_sector (_MountInfo.rootOffset + sector);

    //! get directory info
    directory = (PDIRECTORY) buf;
```

ルートディレクトリから読み込んでいます。ルートクラスタは、ファイルシステムがマウントされたときにBios Parameter Block (BPB) から取得した情報を含む_MountInfoに格納されています。_MountInfo.rootOffsetには、ルートディレクトリの最初のクラスタが格納されています。ルート

ディレクトリには、最大で 224 個の DIRECTORY エントリが含まれます。1つの DIRECTORY エントリーは 32 バイトで、 $224 \times 32 = 7168$ バイト、 $7168 \text{ バイト} \div 512 \text{ バイト} (512 \text{ バイトで } 1 \text{ クラスター}) = 14$ となります。つまり、ルートディレクトリは 14 個のクラスターで構成されています。これを知っていれば、ディレクトリ全体を一度に読み込むのではなく、セクターごとに読み込み、各部分を解析することができます。

```
///! 16 entries per sector
for (int i=0; i<16; i++) {
    ///! get current filename
    char name[11];
```

```

    memcpy (name, directory->Filename, 11);
    name[11]=0;

    //! find a match?
    if (strcmp (DosFileName, name) == 0) {

```

1つのDIRECTORYエントリーが32バイトであることを知ると、1クラスタ512バイト ÷ 32バイト = 16。つまり、1つのセクタに16個の DIRECTORYエントリがあることになります。そこで、各エントリをループしてファイル名を比較し、探しているファイルやディレクトリを見つけます。file.currentClusterには、後で読むためのファイルの最初のクラスタが格納され、file.fileLengthには、ファイルのサイズがバイト単位で格納されます。directory->Attribには、ファイルの属性が格納されます。ここでは、その DIRECTORYエントリー属性に基づいて設定しています。

```

        //! found it, set up file info
        strcpy (file.name, DirectoryName);
        file.id          = 0;
        file.currentCluster = directory->FirstCluster;
        file.eof          = 0;
        file.fileLength   = directory->FileSize;

        //! set file type
        if (directory->Attrib == 0x10)
            file.flags = FS_DIRECTORY;
        else
            file.flags = FS_FILE;

        //! return file
        return file;
    }

```

あと少し...。ファイルやディレクトリがまだ見つからない場合は、次の DIRECTORYエントリーに移動します。ファイルが見つからない場合は、FS_INVALIDを設定して戻ります。

```

        //! go to next directory
        directory++;
    }
}

```

```

//! unable to find file
file.flags = FS_INVALID;
return file;
}

```

これで完了です。上記のルーチンはFAT12のディレクトリやファイルに対して動作します。このルーチンを呼び出すと、ルート・ディレクトリ内の任意のフォルダやファイル名を検索し、その情報を返します。

サブディレクトリ

古いバージョンのFAT12はフラットでしたが、このファイルシステムの新しいバージョンではサブディレクトリをサポートしています。これにより、ディレクトリを利用して多くのファイルをより簡単に管理できるようになった。例えば、大規模なOSでは、OS固有のファイルをシステム・ディレクトリに分けたり、ユーザー・プロファイルを含むユーザー・ディレクトリに分けたりするのが良いでしょう。

サブディレクトリとは、普通のファイルに DIRECTORYフラグを設定したものです。このため、まずはファイルの読み方を知る必要があります。

ファイル読み込み

フォーマット

さて、これでディレクトリを解析してファイルを探すことができるようになりました。次に、ファイルの内容を読み取る方法が必要です。技術的には、ファイルのディレクトリエントリ構造のFirstClusterフィールドを指定するだけで、ファイルの最初の512バイトを読み取ることができることを覚えておいてください。1つ以上のクラスタを読み取るには、ファイルアロケーションテーブル（FAT）を解析する必要がある。

FATはクラスタ番号を含むいくつかのエントリで構成されていることを思い出してください。これらのエントリのサイズはファイルシステムに依存します。FAT12はエントリあたり12ビット、FAT16は16ビット、FAT32は32ビットです。

FATはリンクされたリストではなく、物理的なディスク全体を表すエントリのテーブルであると考える。ディスクの最初のクラスタは、FATの最初のエントリで表されます。ディスクの最初のクラスタはFATの最初のエントリで表現され、2番目のクラスタは2番目のエントリで表現され、以下同様です。つまり、クラスタとFATエントリの間には1対1の関係があります。これにより、FAT12でのファイルの読み書きが容易になります。

ファイルの読み込み

ファイルを読み取るには、そのファイルの現在のクラスタを読み取るだけです。その後、FATテーブルを解析してディスク上の次のクラスタを探します。次のクラスターを見つけたら、次のファイルを読むために「現在のクラスター」を更新します。

読み込むクラスタは、ファイルのオープン時に設定されました。このルーチンの最初の呼び出しでは、file->currentClusterは

DIRECTORY->FirstClusterです。

このクラスタは、ディスク上のデータ領域へのオフセットです。FAT12フォーマットのディスクのフォーマットを思い出して、FATとデータ領域の位置を確認してみましょう。



Boot Sector	Extra Reserved Sectors	File Allocation Table 1	File Allocation Table 2	Root Directory (FAT12/FAT16 Only)	Data Region containing files and directories.
-------------	------------------------	-------------------------	-------------------------	-----------------------------------	---

各FATは9個のセクタを取ることを覚えておいてください。2つのFATがあるので、 $9+9=18$ となります。また、前節でルートディレクトリが14セクタであると結論づけました。 $18+14=32$ 。これは、FATとルートディレクトリの両方が占めるセクタ数です。ここまで計算式は、 $32 + \text{file-} > \text{currentCluster}$ です。1を引く必要があり、 $32 + (\text{file-} > \text{currentCluster} - 1)$ となります。これが読み込まれるセクタであり、ファイルデータを含みます。

```
void fsysFatRead(PFILE file, unsigned char* Buffer, unsigned int Length)
{
    if (file) {
        //! starting physical sector
        unsigned int physSector = 32 + (file->currentCluster - 1);
        //! read in sector
        unsigned char* sector = (unsigned char*) flpydsk_read_sector ( physSector );
        //! copy block of memory
        memcpy (Buffer, sector, 512);
    }
}
```

次のクラスタを読み込むためには、FATテーブルを解析する必要があります。FATテーブルは9セクタあるので、9セクタすべてを読むのではなく、どのセクタを読むべきかを決定する。

まず、次のクラスターがどこにあるのか、バイトオフセットを取得します。これを行うためには、クラスタの値にクラスタのサイズを掛けます。これはFAT_Offsetに格納されます。FAT32クラスタのサイズは4バイトなので、FAT32を使用している場合は4倍します。FAT16を使用している場合は、クラスタ・エントリごとに2バイトを使用するので、2倍します。しかし、FAT12ではどうでしょうか？FAT12はクラスタエントリあたり12ビットを使用する。これは8ビット(1バイト目)と4ビット(2バイト目)。4ビットは8ビットの半分なので0.5)なので、クラスタエントリあたり1.5ビットになります。

この後、このバイト・オフセットをセクタのサイズで割るだけで、読み込むべきFATのセクタが得られる。リマンダーはこのセクタ内のオフセットで、FATから読み出すためのクラスタとなります。これがentryOffsetになります。

FATはuint8_t FAT [SECTOR_SIZE*2]と定義されています。FATのセクタを1つではなく2つメモリに読み込んでいることに注目してください。なぜこのようなことをするのでしょうか？セクタ・サイズが512バイトであることを知ると、 $512\text{バイト} \times 8 = 4096\text{ビット}/\text{セクタ}$ となります。 $4096\text{ビット} \div 12\text{ビット}(\text{FATエントリの場合})$ で $341.3333\dots$ などとなります。つまり、あるエントリは第1セクタと第2セクタの間に位置することになります。これでは、ファイルを読み込むときに問題が生じます。このため、第1セクタの最後のクラスタ値が破損しないように、追加のセクタをロードする必要があります。

```
unsigned int FAT_Offset = file->currentCluster + (file->currentCluster / 2); //multiply by 1.5
unsigned int FAT_Sector = 1 + (FAT_Offset / SECTOR_SIZE);
unsigned int entryOffset = FAT_Offset % SECTOR_SIZE;

//! read 1st FAT sector
sector = (unsigned char*) flpydsk_read_sector ( FAT_Sector );
memcpy (FAT, sector, 512);

//! read 2nd FAT sector
sector = (unsigned char*) flpydsk_read_sector ( FAT_Sector + 1 );
memcpy (FAT + SECTOR_SIZE, sector, 512);
```

FATセクターが読み込まれた後、クラスタ番号を読み込んでいく。

ここで問題が発生します。8ビットの値を読み取る場合、クラスタ値の12ビット全体を読み取ることはできません。そこで、uint16_tを使って16ビットを読み取ることにしました。もちろん、今度は12ビットの値のビット数が多すぎるという問題があります。

もう少し詳しく見てみましょう。これが私たちのFATだとします。ここではFATをバイトに分割し、12ビットのエントリをマークします。(これは第6章からの引用です。)

Note: Binary numbers separated in bytes.
Each 12 bit FAT cluster entry is displayed.

01011101	0111010	01110101	00111101	00111101	0111010	0011110	0011110
-0 cluster	---	1st cluster	---	2nd cluster---	3rd cluster-	4th cluster ---	---

偶数クラスタは第1バイトのすべてを、第2バイトの一部を占めていることに注目してください。また、すべての奇数クラスタは、第1バイトの一部をオコピしますが、第2バイトのすべてをオコピすることに気がつきます。

このことを念頭に置くと、クラスタが偶数の場合は、次のクラスタに属するため、上位4ビットをマスクアウトする。クラスターが奇数の場合は、4ビット下にシフトします(最初のクラスターで使用したビットを破棄するため)。

さて、以上のことを踏まえた上で、この関数を完成させましょう。

```
///! read entry for next cluster
uint16_t nextCluster = *( uint16_t*) &FAT [entryOffset];

///! test if entry is odd or even
if( file->currentCluster & 0x0001 )
    nextCluster >>= 4;           //grab high 12 bits
else
```

```

        nextCluster &= 0x0FFF; //grab low 12 bits
    //!
    if (nextCluster >= 0xff8) {
        file->eof = 1;
        return;
    }
    //!
    if (nextCluster == 0) {
        file->eof = 1;
        return;
    }
    //!
    file->currentCluster = nextCluster;
}

```

ファイルの書き込み

(誠人の声) チャプターアップデートで完成するんだ

サブディレクトリ

サブディレクトリとは、DIRECTORY属性が設定されたファイルのことです。サブディレクトリから読み取るためには、そのディレクトリ名を持つディスク上のFAT12ファイルを探し出し、FATを使った他のファイルと同様の方法で読み取ればよいのです。

ファイルが読み込まれた後、最初のバイトから最後のバイトまでは、単なるDIRECTORYエントリーの配列です。このディレクトリを読むためには、ルートディレクトリと同じようにDIRECTORYエントリーを解析します:-) これがディレクトリ内のファイルやフォルダになります。見てみましょう。

```

FILE fsysFatOpenSubDir (FILE kFile,
                        const char* filename) {

    FILE file;

    //!
    //! get 8.3 directory name
    char DosFileName[11];
    ToDosFileName (filename, DosFileName, 11);
    DosFileName[11]=0;
}

```

filename は検索したいファイルまたはディレクトリ、kFile は解析したいサブディレクトリです。 myfile.txt のような入力ファイル名を、DOS 8.3ファイルシステムのフォーマットである「MYFILE TXT」に変換し、DosFileNameに格納します。

```

//! read directory
while (! kFile.eof) {

    //!
    //! read directory
    unsigned char buf[512];
    fsysFatRead (&file, buf, 512);

    //!
    //! set directort
    PDIRECTORY pkDir = (PDIRECTORY) buf;
}

```

ファイルは、解析したいサブディレクトリです。FAT12では普通のファイルなので、ファイルのセクタを読み込むことになります。ファイルは DIRECTORYエントリの配列で構成されています。 DIRECTORYのメンバーに簡単にアクセスできるように、pkDirを使ってセクターの内容を指示します。 それでは、ディレクトリを検索してみましょう...

```

//! 16 entries in buffer
for (unsigned int i = 0; i < 16; i++) {

    //!
    //! get current filename
    char name[11];
    memcpy (name, pkDir->Filename, 11);
    name[11]=0;

    //!
    //! match?
    if (strcmp (name, DosFileName) == 0) {

```

DIRECTORYの各エントリは32バイトです。セクター(FAT12のクラスタでもある)は512バイトです。 512バイト ÷ 32バイト = 16個の DIRECTORY エントリがセクタごとに存在します。そこで、16個のエントリーすべてをループさせて名前を比較します。検索しているファイル名と一致するファイル名が見つかれば、そのファイルは見つかったことになります。

```

        //!
        //! found it, set up file info
        strcpy (file.name, filename);
        file.id          = 0;
        file.currentCluster = pkDir->FirstCluster;
        file.fileLength   = pkDir->FileSize;
        file.eof          = 0;

```

```

    //! set file type
    if (pkDir->Attrib == 0x10)
        file.flags = FS_DIRECTORY;
    else
        file.flags = FS_FILE;
    //! return file
    return file;
}

```

ファイルが見つかると、FILE構造を記入します。最初のファイルクラスタ（後で読めるように）、ファイルサイズ（EOFがいつなのかわかるよう）に、属性（ファイルまたはディレクトリ）です。

ファイルが見つからなかった場合は、次のエントリに進むだけです。このループは、ファイルの最後まで続きます。ファイルが見つからなかった場合は、FS_INVALIDを設定して戻ります。

```

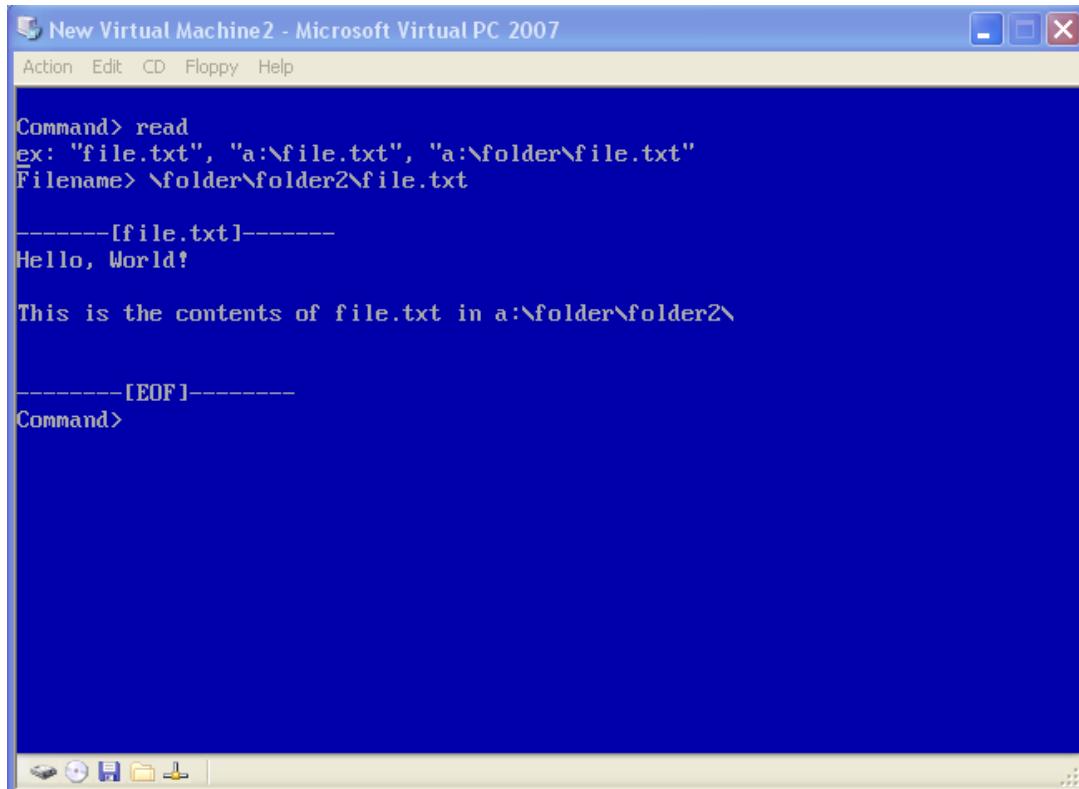
        //! go to next entry
        pkDir++;
    }

    //! unable to find file
    file.flags = FS_INVALID;
    return file;
}

```

このルーチンと私たちの FsysFatDirectory ルーチンの類似点に注目してください。

Demo



OSでファイルを見る

デモダウンロード

本章のデモでは、これまでに説明した内容をすべて盛り込み、VFSとFAT12のミニドライバを実装しています。複数のファイルシステムのサポート、

ディスク・デバイスのサポート、サブディレクトリのサポート、ファイルの読み込みと表示が可能です。

また、このデモは大きなファイルを表示することができ、マルチクラスターファイルのための「キーを押して続行」機能を実装しています。このデモでは、CRT の string.c に strchr() ISO C ルーチンを組み込み、テキストの解析を支援しています。また、このデモでは read

コマンドを使って、生のセクタではなく、ファイルを探して表示できるようになりました。

このデモでは、ボリュームマネージャは非常にシンプルで、fsys.cpp に実装されています。ファイルシステムの登録と解除、ファイルシステムの抽象化を管理します。volOpenFile() を呼ぶと、ファイルを開くことができます。デフォルトでは a:\file.txt を開きますが、任意のディレクトリの任意のファイルを開くように呼び出しても動作します。

すべてのファイルシステムがサブディレクトリをサポートしているわけではありません。そのため、サブディレクトリのサポートはファイルシス

2021/11/15 13:11

Operating Systems Development Series

テムのドライバに任せています。その代わり、ボリュームマネージャはパス名のドライブレターパートのみを処理します。たとえば、`volOpenFile("a:$folder$file.txt")` を呼び出すと、`volOpenFile` は 「\$folder\$file.txt」 をデバイス「a」に登録されているファイルシステムに渡します。ファイルシステムのドライバーは、ディレクトリのパス名を解析し、サブディレクトリやファイルを開く役割を果たします。

FAT12ミニドライバの場合、この特別なルーチンは`fsysFatOpen()`であり、ディレクトリ・パス(例えば"~/~")を解析し、ファイルやディレクトリを解析して読み取るための他のファイル・システム・ルーチンを呼び出す役割を担っている。

以上です:-) この章がFAT12を扱う最後の章となるでしょう。このため、ファイルやディレクトリのディスクへの書き込みについては、もう少し後にアップデートを予定しています。

結論

この章は楽しかったですね。ディスクからファイルを読み込むことができるようになりました。分かっている、分かっている、"そろそろだな!"と。:) これで、マルチタスクやプログラムの実行に向けて、大きく飛躍することができます。しかし、マルチタスク化の前に、ローダーについて説明しておきましょう。ローダーは、プログラムの読み込みと実行、アドレス空間へのマッピングを行います。また、アドレス空間でのヒープ管理やスタック管理についても説明する必要があります。

メモリ管理の章を大幅に更新する予定なので、ヒープとスタックの管理をメモリ管理の章の次の章に移すかもしれません。いずれにしても、変更点については隨時お知らせしていきます。

しかし、これは、私たちがマルチタスクに飛び込む時期が近づいていることを意味しています。その後は? ユーザーモードです。では、次回をお楽しみに。

~マイク

*BrokenThorn Entertainment*社。現在、DoEとNeptune Operating Systemを開発中です。質問やコメントはありますか？お気軽にお問い合わせください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。



Chapter 21

Home



Chapter 23



オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - ユーザーランド

by Mike, 2010

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

はじめに

Welcome!

前章では、VFSを見て、テキストファイルを読み込んで表示しました。このVFSを使って、実行可能なプログラムファイルも読み込むことができます。これには、ドライバー、プログラムソフト、共有、ランタイムライブラリなどが含まれます。

この章では、ユーザーランド・ソフトウェアのサポートに踏み込んでいきます。また、システムAPIとその仕組みについてもご紹介します。さあ、始めましょう。

保護レベル

アセンブリ言語の輪

カーネルランド

第5章では、アセンブリ言語で使われるリングの概念を簡単に説明しました。このリングは、さまざまな保護レベルを表しています。これらの保護レベルは、ハードウェアの詳細であり、ハードウェアによって実装されます。

リング0で動作するソフトウェアは、最も制御しやすい。例えば、ハードウェアPIO、MMIO、プロセッサのハードウェア制御やテーブル（CPUのキャッシュ制御やMMRなど）などがあります。

特権的な命令のリストは第7章で紹介しましたが、ここでは完璧を期すためにリストアップしています。

保護レベルが0より大きい状態で実行されているソフトウェアが上記の命令を実行しようとした場合、プロセッサは以下を生成します。

保護障害 (#PF) の例外。

Privileged Level Instructions	
Instruction	Description
LGDT	Loads an address of a GDT into GDTR
LLDT	Loads an address of a LDT into LDTR
LTR	Loads a Task Register into TR
MOV Control Register	Copy data and store in Control Registers
LMSW	Load a new Machine Status WORD
CLTS	Clear Task Switch Flag in Control Register CRO
MOV Debug Register	Copy data and store in debug registers
INVD	Invalidate Cache without writeback
INVLPG	Invalidate TLB Entry
WBINVD	Invalidate Cache with writeback
HLT	Halt Processor
RDMSR	Read Model Specific Registers (MSR)
WRMSR	Write Model Specific Registers (MSR)
RDPMC	Read Performance Monitoring Counter
RDTSC	Read time Stamp Counter

このため、リング0で動作するソフトウェアをカーネルランドまたはカーネルモードと呼びます。リング0はスーパーバイザーモードとも呼ばれます。

これまでのシリーズで書いてきたソフトウェアは、すべてカーネルモードのソフトウェア、つまりカーネルとミニドライバでした。マイクロカーネルやハイブリッドでは、通常、シリーズで使用しているものよりも高度なドライバー・インターフェーシング・スキームを採用しており、ドライバーを適切にインストールしたり、ドライバーをカーネルとは完全に分離したユーザー・モードで実行したりすることができます。カーネルの一部をユーザー・モードにすることも可能ですが、それは設計次第です。

ユーザーランド

これは、マシンを保護するためのもので、リング1～3で動作するソフトウェアに起因するエラーが発生した場合、プロセッサは一般保護（#GP）例外を使用してシステムエグゼクティブまたはカーネルに問題を通知します。

ほとんどのオペレーティングシステムは、カーネルモードとユーザーモードの2つのモードシステムを採用しています。x86ファミリーでは4つの保護モードをサポートしていますが、これらのオペレーティングシステムでは、アーキテクチャ間の移植を容易にするために2つのモードのみを使用しています。

これらのOSでは、カーネルモードのソフトウェアはリング0で動作し、ユーザーランドのソフトウェアはリング3で動作するように設計されています。リング1と2は使用されません。ドライバー・ソフトウェアは、ハードウェア・デバイスにアクセスするためにリング0で動作するか、提供されているドライバーAPIまたはシステムAPIを使用してハードウェア・デバイスと通信するためにリング3で動作します。

ユーザーモードのソフトウェアは、ハードウェアデバイスに直接アクセスすることができないため、システムのタスクを完了するためにオペレーティングシステムに通知する必要があります。これは、文字の表示、ユーザーからの入力、文書の印刷などを含みます。これらの機能は、ライブラリやAPIの形でユーザーモードソフトウェアに提供される。これらのライブラリやAPIは、システムAPIと通信します。

システムAPI.....この言葉を目にしたことがある人は多いだろう。システムAPIについては、もう少し詳しく見てみましょう。今回は、ユーザーモードについて詳しく見ていきましょう。

リング -1

最近のプロセッサには、ハイパーバイザのリング0アクセスを許可する特別な保護レベルを持つものがあります。これは「リング」と呼ばれることがあります。

-1".

ユーザーランドへようこそ

ユーザーモードに入るにはいくつかのステップがあります。(さあ、簡単だとは思わなかったでしょう。) でも、それほど悪いことではありません。

ステップ1：グローバル記述子テーブル

まず、グローバルディスクリプター・テーブル(GDT)に戻る必要があります。GDTは、プロテクトモードを初めて設定する際に必要となった大きな醜い構造体です。GDTには、プロセッサの情報を含む8バイトのエントリのリストが含まれています。GDTのエントリのビットフォーマットをもう一度見てみましょう。(重要な部分は太字にしました)

56~63ビット目 ベースアドレスの24~32ビット目

ビット55 : グラニュラリティ

0: なし

1: リミットが4K倍になります。

ビット54 : セグメントタイ

プ

0: 16ビット

1: 32ビット

ビット53 : 予約済み-0にすべき

ビット52 : OS用の予約

ビット48~51 : セグメントリミットのビット16~19

ビット47 : セグメントがメモリ内にある (仮想メモリで使用)

ビット45-46 記述子の特権レベル 0: (Ring

0) Highest

1: (Ring 1)

2: (Ring 2)

3: (リング3) 最下位ビット

44 : ディスクリプタービット

0 : システム記述子

1: コードまたはデータ記述子

ビット41-43 ディスクリプタ・タイプ

ビット43 : 実行可能セグメント

0 : データセグメント

1: コードセグメント

ビット42 : 拡張方向 (データセグメント) 、準拠 (コードセグメント)

ビット41 : 読み出し可能、書き込み可能

0 : 読み出しのみ (データセグメント) 、実行のみ (コードセグメント)

1 : 読み取りと書き込み (データセグメント) 、読み取りと実行 (コードセグメント)

ビット40 : アクセスピット (仮想記憶で使用) ビット

16~39 ビット16-39 : ベースアドレスの0-23ビット ビ

ット0-15 : セグメントリミットの0-15ビット

そうですか。上記のDPL (Descriptor Privilege Level) ビットは、そのディスクリプターに使用される特権レベルを表しています。つまり、これらのビットを3に設定することで、その記述子を実質的にユーザーモードの記述子することができます。

そこでまず、GDTに2つの新しいディスクリプターを作成します。1つはユーザーモードデータ用、もう1つはユーザーモードコード用です。これは、i86_gdt_initializeを変更して、ユーザーモードコードとデータのための2つの新しいGDTエントリを追加することで行われます。それでは早速やってみましょう。

```
///! initialize_gdt
int i86_gdt_initialize () {

    ///! etc...

    ///! set default user mode code descriptor
    gdt_set_descriptor (3,0,0xffffffff,
        I86_GDT_DESC_READWRITE|I86_GDT_DESC_EXEC_CODE|I86_GDT_DESC_CODEDATA|
        I86_GDT_DESC_MEMORY|I86_GDT_DESC_DPL,
        I86_GDT_GRAND_4K | I86_GDT_GRAND_32BIT | I86_GDT_GRAND_LIMITHI_MASK);

    ///! set default user mode data descriptor
    gdt_set_descriptor (4,0,0xffffffff,
        I86_GDT_DESC_READWRITE|I86_GDT_DESC_CODEDATA|I86_GDT_DESC_MEMORY|
        I86_GDT_DESC_DPL,
```

```
I86_GDT_GRAND_4K | I86_GDT_GRAND_32BIT | I86_GDT_GRAND_LIMITHI_MASK);

// etc...

return 0;
}
```

上記のコードは、他のGDTエントリを作成するときに行ったものと同じですが、1つだけ変更があります。I86_GDT_DESC_DPLフラグに注目してください。これにより、両方のDPLビットが2に設定され、ユーザー mode（リング3）用になります。これらのフラグはすべて、以前の章でプロテクトモードについて説明したときのものです。

必要なのは、これだけです。なお、ユーザー mode コード記述子は GDT のインデックス 3 に、ユーザー mode データ記述子はインデックス 4 にインストールされています。セグメントレジスタには、使用するセレクタのオフセットが含まれていることを覚えておいてください。各 GDT エントリは 8 バイトサイズなので、コードセレクタ 0x18 (8*3) 、データセレクタ 0x20 (8*4) となります。

したがって、これらのセレクタを使用するには、上記のセグメントセレクタの 1 つを、使用するセグメントレジスタにコピーするだけです。

DPL

DPL (Descriptor Protection Level) とは、セグメント記述子の保護レベルのことです。例えば、カーネルのコードセグメントとデータセグメントの DPL は、リング 0 のアクセスに対して 0 となります。

RPL

Requested Protection Level (RPL) は、ソフトウェアが CPL をオーバーライドして新しい保護レベルを選択できるようにするものです。これにより、ソフトウェアは、リング 0 から リング 3 など、他の保護レベルの変更を要求することができます。RPL はディスクリプターセレクターのビット 0 と 1 に格納されています。

待って、何？セグメントセレクタは、GDT への単なるオフセットであることを覚えておいてください。例えば、0x8 バイトは リング 0 のコード記述子のオフセットです。0x10 は、データセレクタのオフセットでした。0x8 と 0x10 はセグメントセレクターです。GDT エントリーはすべて 8 バイトなので、セグメントセレクターの値は常に 8、16、24、32 など、8 の倍数になります。8 は 2 進法では 1000 です。つまり、セグメントセレクターの値がどのようなものであっても、下位 3 ビットはゼロになります。

RPL は、セグメントセレクタの下位 2 ビットに格納されます。セグメントセレクターが 0x8 の場合、RPL は 0 になります。0xb (0x8 だが、最初の 2 ビットがセットされており、1000 ではなく 1011 のバイナリ) の場合、RPL は 3 になります。

CPL

カレントプロテクションレベル (CPL) は、現在実行中のプログラムの保護レベルです。CPL は SS と CS のビット 0 と 1 に格納されています。

GDT エントリのサイズは 8 バイトであることを覚えておいてください。プロテクトモードのセグメントレジスタには、セグメントセレクタ (GDT エントリのオフセット) が含まれているため、下位 3 ビットはゼロであることが保証されています。CS と SS の下位 2 ビットは、ソフトウェアの CPL を格納するために使用されます。

保護レベル

ソフトウェアがセグメント・レジスタに新しいセグメントをロードしようとすると、プロセッサはソフトウェアの CPL と ロードしようとしているセグメントの RPL とのチェックを行います。RPL が CPL よりも高ければ、ソフトウェアはセグメントをロードできます。そうでない場合、プロセッサは一般保護フォルト (#GP) を発生させます。

RPL の仕組みを理解しておくことは、ユーザー mode に切り替える際に必要な情報となります。

ステップ2：スイッチ

これで、ユーザー mode に切り替えることができます。

ジャンプを実行するには 2 つの方法があります。SYSEXIT 命令を使う方法と、IRET を使う方法です。どちらの方法にも一長一短がありますので、詳しく見ていきましょう。本連載では、移植性を考慮して IRET を使用しています。

SYSEXIT命令

このセクションは、今後も発展させていく予定です。

IRET / RETD 命令

SYSEXIT を使うよりも移植性が高いため、多くの OS がこの方法を採用しています。より大きな OS では、SYSEXIT が使えない場合のバックアップ方法として、この方法をサポートしているかもしれません。

では、IRET はどのようにして切り替えを行うのでしょうか。第 3 章で説明した、モードを切り替えるときの方法を思い出してください。IRET は ラップリターン命令です。IRET を実行すると、ユーザー mode のコードに戻るよう、スタックフレームを調整することができます。

IRETD が実行されると、スタックには以下のものがあると期待されます。

- SS
- ESP
- EFLAGS
- CS
- EIP

IRET_Dにより、プロセッサはスタックから取得したCS:EIPにジャンプします。また、スタックから上記の値をEFLAGSレジスタに設定します。SS:ESPには、スタックから取得したSSとESPの値が設定されます。

これらの値は、INT命令が実行されると自動的にスタックにプッシュされます。このため、通常の場合、これらの値は変更されません。しかし、これらの値を変更することで、IRETにモードスイッチを実行させることができます。

さて、まずはセグメントセレクターの設定です。下位2ビットが希望するRPLを表すことを思い出してください。ここでは、ユーザーモードに3を設定します。では、その設定を行います。

```
void enter_usermode () {
    _asm {
        cli
        mov ax, 0x23      ; user mode data selector is 0x20 (GDT entry 3). Also sets RPL to 3
        mov ds, ax
        mov es, ax
        mov fs, ax
        mov gs, ax
    }
}
```

ここで、ユーザーモードへの切り替えを行います。これは、IRET用のスタックフレームを構築し、IRETを発行することで行われます。

```
        push 0x23          ; SS, notice it uses same selector as above
        push esp           ; ESP
        pushfd            ; EFLAGS
        push 0x1b          ; CS, user mode code selector is 0x18. With RPL 3 this is 0x1b
        lea eax, [a]       ; EIP first
        push eax

        iretd
a:
        add esp, 4 // fix stack
    }
}
```

スタックフレームが上のリストにあったものと一致していることに注目してください。IRETD命令により、上記のコードではリング3内で0x1B:aが呼び出されます。

しかし、ちょっとした問題があります。上記のルーチンを使用したり、カーネル内で別の方法でユーザーモードに切り替えようすると、ページフォルト (PF) の例外が発生します。これは、カーネルのページがカーネルモードアクセス専用にマッピングされているためです。この問題を解決するには、別の方法でユーザーモードに移行するか、ユーザーモードのソフトウェアがアクセスできるようにカーネルをマッピングする必要があります。今のところ、ユーザーモードのソフトウェアがアクセスできるようにカーネルをマッピングするだけです。これにはvmmngr_initialize()の更新が必要です。

ルーチンで、PTEとPDEのUSERビットを設定します。

より複雑なオペレーティングシステムでは、このアプローチは使われません。この方法は、ユーザーモードのソフトウェアがアクセスできるようにカーネルページをマッピングした場合にのみ機能しますが、これは良くありません。より推奨されるアプローチは、カーネルだけのアクセスのためにカーネルページをマップしておき、ユーザプログラムをロードするときに、カーネルのローダコンポーネントがユーザモードページをマップするようにすることです。スタックとヒープのアロケータは、プログラムのスタックとヒープの領域をユーザモードにマッピングします。この現在の方法では、カーネルのスタックをユーザーランドと共有しています。大きなシステムでは、このような方法をとるべきではありません。

v8086モードへの移行

v8086モードに入るためには、ユーザーモードのタスクが必要です。したがって、上記の手順を実行すれば、v86モードにも入ることができます。ただし、1つだけ若干の修正が必要です。

EFLAGSレジスタのフォーマットを思い出してください。ビット17 (VM) は、v8086モード制御フラグです。IRETを実行する際にEFLAGSの値をスタックにプッシュするので、v86モードに入るためには、EFLAGSのビット17をセットしてからスタックにプッシュすればよい。これにより、IRETはリターン時にEFLAGSレジスタのVMビットを設定します。

これだけで、v8086モードに入ることができます。

デザインに関する注意事項

上記の方法は、ユーザーモードに入るための簡単な方法ですが、コストがかかります。上記の方法が機能するためには、リング3ソフトウェアがカーネルメモリにアクセスできるようにカーネル領域をマッピングする必要があります。このため、リング3で実行中のソフトウェアは、プロトコトモードによる制限はあるものの、カーネルルーチンを直接呼び出したり、カーネル空間をゴミ箱に入れたりすることができます。

上記の問題を解決する方法として、カーネルメモリをリング0ソフトウェア用に予約しておくことが考えられます。カーネルのローダコンポーネントは、プログラムをロードしている間に、プロセスに必要なリング3のメモリ領域をマッピングすることができます。

この点については、次章でOS用のローダーを開発する際に詳しく説明します。

カーネルランドへの切り替え

ステップ1：TSSの設定

x86アーキテクチャは、ハードウェアによるタスク切り替えをサポートしています。つまり、プロセッサが異なるタスクを選択できるように、ハードウェアで定義された構造がアーキテクチャに含まれています。

最近のOSの多くは、移植性を考慮して、ハードウェアのタスクスイッチを利用ていません。これらのOSでは、一般的にソフトウェアによるタスクスイッチ方式が採用されています。

タスクステートセグメント (TSS)

TSSの構造はかなり大きい。

```
#ifdef _MSC_VER
#pragma pack (push, 1)
#endif

struct tss_entry {
    uint32_t prevTss;
    uint32_t esp0;
    uint32_t ss0;
    uint32_t esp1;
    uint32_t ss1;
    uint32_t esp2;
    uint32_t ss2;
    uint32_t cr3;
    uint32_t eip;
    uint32_t eflags;
    uint32_t eax;
    uint32_t ecx;
    uint32_t edx;
    uint32_t ebx;
    uint32_t esp;
    uint32_t ebp;
    uint32_t esi;
    uint32_t edi;
    uint32_t es;
    uint32_t cs;
    uint32_t ss;
    uint32_t ds;
    uint32_t fs;
    uint32_t gs;
    uint32_t ldt;
    uint16_t trap;
    uint16_t iomap;
};

#endif
#pragma pack (pop, 1)
#endif
```

TSSは、ハードウェアによるタスクスイッチが行われる前のマシンの状態に関する情報を格納するために使用されます。たくさんのメンバーがありますので、見てみましょう。

一般的な分野。

タスク切り替え前のLDT,EIP,EFLAGS,CS,DS,ES,FS,GS,SS,EAX,EBX,ECX,EDX,ESP,EBP,ESI,EDIの状態

prevTSS - タスクリストの前のTSSのセグメントセレクタ

cr3 - PDBR、現在のタスクトラップのページディレクトリのアド

レス

Bit 0: 0: 無効, 1: タスクからタスクへの切り替え時にデバッグ例外を発生 iomap - TSSベースからI/Oパーティ

ションおよび割り込みリダイレクションピットマップへの16ビットオフセット esp0,esp1,esp2 - リング

0,1,2のESPスタックポインター

ss0,ss1,ss2 - リング0,1,2のSSスタックセグメント

これらのフィールドのほとんどは非常に単純です。このため、この構造体のいくつかのフィールド、特にリング0スタックとセレクタフィールドを設定する必要があります。

Step 2: TSSのインストール

記述子セグメント

TSSはその名前からもわかるようにセグメントです。他のセグメントと同様に、TSSにもGDTへのエントリが必要です。これにより、TSSを制御することができます。タスクがビジーかインアクティブかの設定、どのソフトウェアがアクセスできるか (DPL) 、その他のフラグを記述子で設定することができます。

Base Address フィールドには、弊社が設定したTSS構造のベースアドレスを入力してください。

LTR命令

LTR (Load Task Register) 命令は、TSRレジスタにTSSをロードするための命令です。例えば、以下のようになります。

```
void flush_tss (uint16_t sel) {
    __asm ltr [sel]
}
```

axは、TSSのセグメントセレクタです。このアーキテクチャは、ハードウェアによるタスク切り替えをサポートしているため、TSRには、現在のタスクを定義するTSSのアドレスが格納されています。

タスクステートレジスタ (TSR) は、TSSセレクタ、TSSベースアドレス、TSSリミットを格納するレジスタです。ただし、ソフトウェアで変更できるのは、TSSセレクタのみです。

TSSのインストール

TSS構造をインストールするには、まずTSSのGDTエントリをインストールします。を呼び出して、TSSを現在のタスクとして選択します。上記のflush_tss。

```
void install_tss (uint32_t idx, uint16_t kernelSS, uint16_t kernelESP) {
    //! install TSS descriptor
    uint32_t base = (uint32_t) &TSS;
    gdt_set_descriptor (idx, base, base + sizeof (tss_entry),
        I86_GDT_DESC_ACCESS|I86_GDT_DESC_EXEC_CODE|I86_GDT_DESC_DPL|I86_GDT_DESC_MEMORY,
        0);

    //! initialize TSS
    memset ((void*) &TSS, 0, sizeof (tss_entry));

    TSS.ss0 = kernelSS;
    TSS.esp0 = kernelESP;

    TSS.cs=0x0b;
    TSS.ss = 0x13;
    TSS.es = 0x13;
    TSS.ds = 0x13;
    TSS.fs = 0x13;
    TSS.gs = 0x13;

    //! flush tss
    flush_tss (idx * sizeof (gdt_descriptor));
}
```

上記のコードでは、TSSはtss_entry構造体のグローバル構造体定義となっています。TSSsのセレクタエントリを、前のタスク（ユーザモードセレクタ）とリング0スタック（カーネルスタック、kernelSS:kernelESPに位置する）に合わせて設定します。flush_tssはTSSをTSRにインストールします。

追加指示

この他にも便利な命令がいくつかあります。これらの命令はすべて、ユーザーモードのソフトウェアで実行できます。

VERR命令

VERR (Verify Segment is Readable) は、セグメントが読み取り可能かどうかを確認するために使用できます。読み取り可能であれば、プロセッサはゼロ・フラグ (ZF) を1に設定します。この命令はどのプロビジョンスペックでも実行できます。

```
verr [ebx]
jz .readable
```

VERW命令

VERW (Verify Segment is Writable) は、セグメントが書き込み可能かどうかを確認するために使用できます。書き込みが可能であれば、プロセッサはゼロ・フラグ (ZF) を1に設定します。この命令はどのプロビジョンスペックでも実行できます。

```
verw [ebx]
jz .readable
```

LSL指導

この命令は、セレクタのセグメントリミットをレジスタにロードするために使用できます。

```
lsl ebx, esp
jz .success
```

ARPL命令

この命令は、セレクタのRPLを調整するために使用することができます。この命令は arpl dest,src という形式で、dest はメモリロケーションまたはレジスタ、src はレジスタです。destのRPLがsrcよりも小さい場合、destのRPLビットはsrcのRPLビットに設定されます。例えば、以下のようにになります。

```
arpl ebx, esp
```

システムAPI

アブストラクト

システムAPIは、ソフトウェアがオペレーティングシステムと対話するためのツール、ドキュメント、およびインターフェースを提供する。オペレーティングシステムによって用語が異なる場合がありますが、基本的な考え方は同じです。例えば、WindowsではこのAPIを「ネイティブAPI」と呼んでいます。

システムAPIは、ソフトウェアがオペレーティングシステムやデバイスドライバと相互作用することを容易にする。システムAPIは、ユーザー モードのソフトウェアとカーネルモードのソフトウェアの間のインターフェースである。ソフトウェアがシステムの情報を必要としたり、ファイルの作成などのシステムタスクを実行する場合、ソフトウェアはシステムコールを呼び出すことになる。

システムコールは、システムサービスとも呼ばれ、オペレーティングシステムが提供するサービスのことです。このサービスは通常、機能やルーチンです。ソフトウェアは、システムのタスクを実行するためにシステムコールを呼び出すことができます。

デザイン

sysenter / sysexit

このセクションは、今後も発展させていく予定です。

ソフトウェインタラプト

ほとんどのシステムAPIは、ソフトウェア割り込みを使って実装されています。ソフトウェアはint 0x21のような命令を使ってオペレーティングシステムのサービスを呼び出すことができます。例えば、DOSのTerminate関数を呼び出すには次のようにします。

```
mov ax, 0x4c00 ; function 0x4c (terminate) return code 0
int 0x21 ; call DOS service
```

上のコードでは、AHは関数番号を含んでいます。int 0x21は、DOSを呼び出すために0x21の割り込みベクターを呼び出します。

上記を動作させるためには、OSは割り込みベクトル0x21のISRをインストールする必要があります。このISRは、AHを比較して正しいカーネルモード関数に制御を渡すFSM (Finity State Machine) となります。以上が、読者の皆様へのデザインです。

ソフトウェア割り込みは、SYSENTERやSYSEXITよりも移植性に優れています。このため、ほとんどのOSがこの方法をサポートしています（他の方法と一緒にサポートしている場合もあります）。このシリーズでは、この方法を使用します。

例

システムAPIは通常、数百のシステムコールで構成されています。

これは、いくつかのオペレーティングシステムと、それらがサポートしている方法の一覧です。INT番号は、上記の方法によるソフトウェア割り込みベクター番号です。

DOS : INT 0x21

Win9x (95,98)の場合。INT 0x2F

WinNT (2k,XP,Vista,7)の場合。INT 0x2E, SYSENTER/SYSEXIT, SYSCALL/SYSRET

211以上の機能

Linuxです。INT 0x80, SYSENTER/SYSEXIT

190以上の機能

基本システムAPI

ステップ1：システムコールテーブル

ほとんどのシステムAPIは、すべてのサービスを含むシステムコールテーブルを実装しています。このテーブルには、スタティック、ダイナミック、自動生成、またはその3つの組み合わせがあります。大規模なオペレーティングシステムでは、通常、システムコールの自動生成されたダイナミックサイズのテーブルを採用しています。これは、このテーブルに含まれる可能性のあるシステムサービスの数が多いためで、手作業で作成するのは非常に面倒です。

今回の目的では、カーネル内にシステムサービステーブルを定義すればよいでしょう。このテーブルには、カーネルの中で呼び出し可能にしたいさまざまな関数のアドレスが含まれています。

```
#define MAX_SYSCALL 3
void* _syscalls[] = {
    DebugPrintf
};
```

うーん、この表はかなり小さいですね。次の章では、このリストにさらに追加していきますが、それほど複雑なものではありません。

DebugPrintfはユーザモードからアクセス可能であり(カーネルページがマッピングされているため)、DebugPrintfはprivedge命令を使用していないため、ユーザモードのソフトウェアは技術的には問題なくこのルーチンを直接呼び出すことができます。オペレーティングシステムや実行ソフトウェアの設計によっては、これがセキュリティや安定性の問題を引き起こす可能性があります。

このような理由から、一般的に、カーネルページはカーネルモードからしかアクセスできないようにしておくことが推奨されています。ソフトウェアに複雑さをもたらしますが、最終的にはその努力に見合う結果が得られるかもしれません。

ステップ2：サービスディスパッチャ

次のステップは、サービスディスパッチャのISRを作ることです。その前に、どのISRを使うかを決めなければなりませんが、ここではLinuxに従って0x80を使います。しかし、多くのOSでは異なるベクターを使用しているので、好きなベクターを使用することができます。では、ISRをインストールしてみましょう。

ISRはHAL層が管理するIDTに格納されていることを思い出してください。また、第15章では、各IDT記述子がそれぞれのDPL設定を持っていることを思い出してください。IDTエンティティーのDPLがCPLよりも小さい場合、GPFが発生します。つまり、ユーザ・モードに入ると、DPL 3のIDTディスクリプタを持つISRしか呼び出せなくなります。リング3のソフトウェアからシステム割り込みを呼び出せるようにしたいので、このISRを正しいフラグでインストールする必要があります。

しかし、現在のHALサブシステムの設計では、setvect()を呼び出すだけでは、特定のフラグを設定することができないため、これを行うことができません。この問題を回避するために、setvect()に第2パラメーターを追加し、オプションのフラグを設定できるようにしました。これはC++のデフォルトパラメータ機能を利用しているため、他のコードを更新する必要はありません。

```
void syscall_init () {
    //! install interrupt handler!
    setvect (0x80, syscall_dispatcher, I86_IDT_DESC_RING3);
}
```

それがすべてです:)

`syscall_dispatcher`は、システムコール用のISRです。このISRは、`_syscalls`で関数を検索して、どのシステムサービスを呼び出すかを決定する必要があります。通常、システムAPIはEAXを使って関数番号を識別します。ここではそれと同じことをするつもりです。上で定義したシステムサービステーブルのおかげで、EAXをインデックスとして使うことができます。そのため、呼び出す関数は`_syscalls[eax]`となります。

```
_declspec(naked)
void syscall_dispatcher () {
    static int idx=0;
    _asm mov [idx], eax

    //! bounds check
    if (idx>=MAX_SYSCALL)
        _asm iretd

    //! get service
    static void* fnct = _syscalls[idx];
```

さて、これで呼び出すべき関数へのポインタができました。しかし、ここでちょっとした問題があります。上記は、EAXで与えられた値に基づいて、必要なサービス関数へのポインタを効果的に取得します。しかし、それがどんな関数なのかわかりません。また、その関数に何を渡せばいいのか、どれだけのパラメタを持っているのかもわかりません。

解決策としては、関数呼び出しのためにすべてのレジスタをスタックにプッシュすることが考えられます。サービスはすべてC言語のルーチンなので、C言語の関数が期待する方法でパラメタを渡さなければなりません。

```
//! and call service
_asm {
    push edi
    push esi
    push edx
    push ecx
    push ebx
    call fnct
    add esp, 20
    iretd
}
```

これで終わりです。) `add esp, 20`は、プッシュした20バイトをスタックからポップします。また、ISRからIRETDで戻っていることに注意してください。の指示を受けています。

システムソフトウェアやエグゼクティブがそれぞれの割込みベクタにISRをインストールした後、ソフトウェアはソフトウェア割込みを発行することでISRを呼び出すことができます。例えば、`syscall_init`を呼び出してISRをインストールした場合、次のようにシステムサービスを呼び出すことができます。

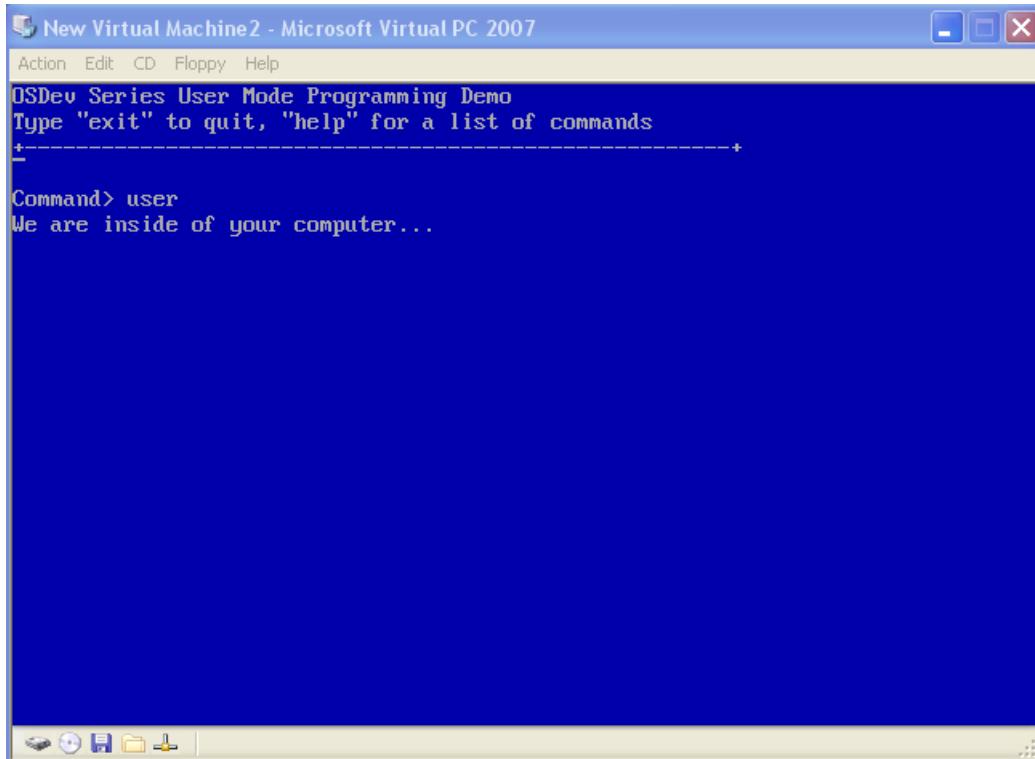
```
_asm {
    xor eax, eax ; function 0, DebugPrintf
    lea ebx, [stringToPrint]
    int 0x80 ; call OS
}
```

デザインに関する注意事項

ほとんどのOSでは、割り込みベクター番号やレジスタの詳細をC言語のインターフェイスに置き換えていました。大規模なOSのシステム・サービスを直接呼び出すことは可能ですが、システムがユーザーランド・ソフトウェアに提供するシステム・サービスを中心に、標準的なCインターフェースを開発することをお勧めします。

大規模なOSでは、ディスプレイにメッセージを表示するシステムサービスを持たないのが普通です。むしろ、ユーザーランドのソフトウェアから呼び出すことができるサービスが含まれており、ユーザーAPIがカーネルモードのサービスやサーバー、デバイスドライバーとやりとりできるようになっていっています。このため、大規模なOSでは、数百の関数呼び出しからなるシステムAPIが一般的です。

Demo



Enter into user mode

Dエモダウンロー

ド

新規・変更ファイル

この章では、シリーズのデモにいくつかのファイルを追加しています。以

下のファイルが含まれます： hal/tss.h

hal/tss.cpp

本章で説明するTSSルーチンを含む hal/user.cpp

ユーザー モード 切り替え ルーチンを含む この章では、以下
のファイルも変更しています。

カーネル/mmngr_virtual.cpp

vmmngr_initialize が更新され、カーネルページをユーザーがアクセスできるようになりました

hal/hal.h/cpp

set_vect() に第2パラメーターを追加しました hal/gdt.h

MAX_DESCRIPTORS は、追加されたGDTエントリ hal/gdt.cppのために6に再定義されました。

た。

ユーザー モード 記述子 のインストールを含むようにアップグレードしました

kernel/main.cpp

新しい変更点を反映して更新

結論

ユーザー ランド へようこそ

これで、ユーザー ランドとカーネル ランドの切り替えに必要なものがすべて揃いました。これで、ユーザー モードのページをマッピングしたり、ユーザー モードでプログラムをロードしたり、実行したりすることができるようになりました。システムがタスクを管理していないため、OS のカーネルにきれいな形で戻る機能はまだありません。これについては次の章で説明します。

次の機会まで。

~マイク

*BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中 質問やコメント
は？お気軽にご連絡ください。*

記事をより良くするために、あなたも貢献してみませんか？よろしければ、ぜひご連絡ください。



Chapter 22

22章 ホーム 第24章

Home



Chapter 24

第



オペレーティングシステム開発シリーズ

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

"デバッグ"がソフトウェアのバグを取り除く作業であるならば

then programming must be the process of putting them in.

- エドガー・ダイクストラ

24: Process Management

by Mike, 2012, 2013

Introduction

1. この章では、プロセス管理とマルチタスクのトピックについて詳しく説明します。これまでの章では、タスクをサポートする基本的なモノリシック・オペレーティング・システムを構築してきました。これまでのOS開発では、複雑な設計をシンプルにするために、複雑さを犠牲にしていましたが、複雑な設計でもある程度の詳細を提供することができました。本章では、この傾向を継続していきます。最後に紹介するデモは、決してOSを開発する唯一の方法ではありません。本章では、以下の項目について説明します。

2. Processes

3. Threads and tasks

4. Looking inside a process

5. Process management

6. Scheduling

7. Linking

Process management

プロセス管理とは、オペレーティングシステムがプロセス、スレッドを管理し、プロセスが情報を共有できるようにし、プロセスリソースを保護し、システムリソースを要求するプロセスに安全な方法で割り当てるプロセスのことです。オペレーティングシステムの開発者にとって、このプロセス管理は非常に困難な作業であり、設計が非常に複雑になる可能性があります。では、それぞれの機能を詳しく見ていきましょう。

プロセスとスレッド

この章の残りの部分では、主にプロセスとスレッドについて説明します。プロセスの作成とは、実行可能なイメージを読み込み、それを実行するための少なくとも1つの実行パス（スレッド）を作成することです。

プロセス間通信

IPC (Inter-Process Communication) は、多くのオペレーティングシステムで採用されている、プロセス間の通信を可能にする技術です。これは通常、メッセージパッシングによって行われます。プロセスは、他のプロセスにメッセージを送信することをオペレーティングシステムに要求し、オペレーティングシステムは、可能であれば他のプロセスにメッセージを送信し、キューに入れることができます。IPCは、ファイル、パイプ、ソケット、メッセージパッシング、シグナル、セマフォ、共有メモリ、メモリマップドファイルなど、さまざまな方法で実装することができます。オペレーティングシステムは、これらのIPCの方法のいずれかまたはすべてを実装することができます。IPCは、ハイブリッドカーネルやモノリシックカーネルの設計でも多用されていますが、マイクロカーネルの設計では間違いなく最も顕著です。

本章では、主にプロセスとスレッドの生成に焦点を当てているため、IPCについては触れません。IPCについてはもう少し後に説明するかもしれません、おそらく本章の追加となるでしょう。

プロセス保護

複数のプロセス同じアドレス空間にロードすると、両方のプロセスがお互いに読み書きできるという基本的な問題が発生します。この問題を解決するには、プロセスをそれぞれの仮想アドレス空間に読み込み、物理アドレス空間の別々の場所にマッピングするのが簡単です。プロセスはより多くのスレッドの作成を要求することができますが、これはプロセスごとに行われるため、すべてのスレッドはプロセスと同じアドレス空間を共有することになります。

また、プロセス保護は、必要最小限のコントロールでプロセスをマッピングすることでも採用されています。例えば、カーネルランドにいる必要のあるプロセスはカーネルスペースに、カーネルランドにいる必要のないプロセスはユーザースペースに配置します。

この章では、プロセスを作成する際に、これらの両方を利用します。プロセスは、ユーザー空間と独自の仮想アドレス空間にマッピングされます。これは、プロセスがカーネルページにアクセスできないことを意味します（つまり、カーネルのスタックや構造体をゴミ箱に入れることはできません）。

ん）。また、プロセスが他のプロセスをゴミ箱に入れることもできません。

資源分配

リソースの割り当てとは、システムのリソース（ファイルやデバイスハンドルなど）を要求するプロセスに安全に渡す方法のことです。シリーズOSの初期状態では、今のところ気にする必要のある資源はありません。しかし、必要に応じて、割り当てられたプロセスリソースは、通常、プロセス制御ブロック内に格納されます。なぜシステムリソースの割り当てを管理する必要があるのかというと、マルチタスク環境において、2つのプロセスが同時に同じファイルを開いて書き込もうとする場合を考えてみてください。

そこで、以下ではプロセスとスレッドの作成を中心に説明します。まず、プロセスとは何か、何がプロセスを構成するのかを明確に定義することから始めます。

Processes

プロセスとは、メモリ上にあるプログラムのインスタンス、またはプログラムの一部のことです。プロセスは、映画やビデオの再生、ゲームのプレイ、あるいはこの文章を書いているエディタの実行など、複雑なタスクを実行するために、オペレーティングシステムやエグゼクティブによって実行されます。要するに、プロセスはプログラムであると言えますが、プログラムは複数のプロセスを含むことができます。例えば、文字列を表示する基本的なプログラムは、独自のプログラムファイルに組み込まれているかもしれません。プログラムをロードすると、OSやエグゼクティブは他のプログラムファイルをロードすることになります。つまり、プロセスが呼び出して使用する実行コードを含む共有ライブラリをダイナミックにロードすることになります。これらのプログラムファイルはすべて同じプロセスの一部であり、1つのプロセスが複数のプログラムファイルのインスタンス、あるいは複数のインスタンスを持つことができるのです。

プロセスは、エミュレートされた環境またはハードウェア環境において、中央処理装置（CPU）または複数のCPUまたはCPUコアによって実行されます。ハイバースレッディングや並列パイプラインをサポートするCPUでは、異なるプロセスの複数の命令を同時に実行することもできます。つまり、プロセスはシーケンシャルに（1命令ずつ）実行されるのではなく、環境やハードウェアの構成に応じて、さまざまな方法で実行される可能性があります。IA32 CPUファミリーのデフォルトでは、これらの機能は無効になっています。つまり、コンピュータの起動時には、CPUはすべての命令を一度に実行します（ただし、命令キャッシュバッファにキャッシングする場合もあります）。しかし、オペレーティングシステムやエグゼクティブがプロセスに対してこれらの機能を有効にした場合、プロセスやシステムはマルチプロセッサに対応できるように設計しなければなりません。しかし、これは高度なテーマであり、エラーが発生しやすいので、上級編で説明します。いくつかのプロセスは、スレッドやタスクに分割することができます。次はこれらについて見ていきましょう。

Threads and tasks

スレッドは、プロセス内の単一の実行経路として定義できます。例えば、最も基本的な例では、メッセージを表示して戻るだけのプログラムがあります。

```
#include <stdio.h>
int main (int argc, char** argv)
{ printf ("Hello, world!"); return 0;
}
```

この例では、プロセスは1つのスレッドを持っています。スレッドはmain()で始まり、プロセスが終了するとスレッドも終了します。（ただし、実際にはそうではないかもしれないことに注意してください。main()を呼び出すランタイムライブラリにスレッドが含まれている可能性があるからです）。では、マルチスレッドの例を見てみましょう。

```
#include <stdio.h> static
int _notExit = 0;

int thread (void* data)
{ while (_notExit)
    /* 便利なことをする */
}
return 0; /* スレッドが終了する（ランタイムに戻り、Terminate Thread を呼び出す */ }

int main (int argc, char** argv)
{ CreateThread (thread);
printf ("Hello, world!"); return 0;
}
```

この例では、CreateThread がオペレーティングシステムを呼び出し、新しい実行フローとして thread() を設定します。CreateThread()が呼び出された後、スレッド()がオペレーティングシステムまたはエグゼクティブによって呼び出され、スレッド()とメイン()の両方の内部で同時に、どちらかが終了するまで実行が続けられます。すべてのプロセスはスレッドですが、スレッドはプロセスではないことに注意してください。プロセスは単一のスレッドを含むことも、多数のスレッドを含むこともできます。プロセス内のスレッドは同じグローバル変数にアクセスして共有することができますが、コンパイラによってはスレッドローカル変数もサポートしています。

スレッドをサポートするオペレーティングシステムは、マルチスレッドに対応していると言われています。Windows、Linux、Mac OSなどがこれにあたる。タスクはスレッドと同義です。タスクは、オペレーティングシステムやエグゼクティブが実行するタスクを表します。したがって、マルチスレッドに対応しているOSは、事実上マルチタスクに対応していることになります。ただし、すべてのマルチタスクOSがマルチスレッディングに対応しているわけではないことに注意が必要です。

プロセスとは一体何なのか？ここでは、「プロセス」を「プログラムのインスタンス」または「プログラムの一部」と定義しましたが、ここでは「プロセス」の内部をさらに詳しく見てみましょう。

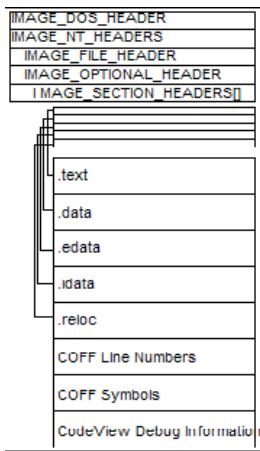
Looking inside a process

プロセスを分解してみると、最も基本的なレベルではコードとデータしかありません。プログラミングの経験がある人なら、これは納得できるだろう。すべてのプログラムは、データに対する動作や操作をCPUに指示する命令に過ぎません。プログラマーが、プログラムの開発を容易にするために、プログラムの「データ」部分と「コード」部分を分離する傾向があるのも理解できるだろう。`.data` と `.text`（プログラムコード用）は、後にプログラムバイナリの中の様々な種類のセクションのうちの2つです。セクションは、プログラムの開発に役立つだけでなく、プログラム・バイナリの中にさまざまな種類のものをどのように格納するかの基準にもなります。

まず、プログラムセクションと、プロセスのアドレス空間にプログラムセクションを配置する方法について説明します。次に
シンボル情報、デバッグ情報、エクスポートとインポートのテーブル、そしてそれらがどのように使用できるか。

プログラムセクション

ここでは、.data と .text という2つのセクション（またはセグメント）について説明しました。プログラムファイルには、これらのセクションとその他のセクションが含まれています。オペレーティングシステムやエグゼクティブは、プロセスが正しく実行されるように、各セクションをアドレス空間にロードすることができます。また、ロード時にセクションを再配置することもできます。これにより、オペレーティングシステムやエグゼクティブは、必要に応じて各セクションの最適な位置を見つけ、それに応じてプロセスを更新することができます。ただし、プログラムファイルのフォーマットによってサポートする内容が異なるため、すべてのプログラムファイルのフォーマットがセクションの再配置をサポートしているわけではありません。



Portable Executable (PE) ファイルフォーマットは、Windows OSで使用される主要なプログラムファイルフォーマットです。PEファイル形式は、コード、データ、リソースデータ、シンボリック情報、マニフェストデータなど、さまざまなセクションをサポートしています。各セクションは、リンクやコンパイラによって書かれたバイナリファイル内に格納される。どのように格納されているかを確認するためには、PEファイル形式のフォーマットを見てみる必要があります。

上記は、PE実行ファイルの中身をイメージしたものです。イメージをメモリにロードし、セクションを再配置する必要がない限り、他のタイプのバイナリファイルと同様にファイルの内容を解析することができます。このようにして、シリーズのブートローダーはカーネルイメージをロードすることができるので、シリーズのカーネルはセクションの再配置を必要としないので、ブートローダはファイルをロードし、ヘッダの中からエントリポイントを見つけて、それを直接呼び出すことができます。これは簡単にできます。

```
/* PEヘッダからエントリポイントを取得 */
IMAGE_DOS_HEADER dosHeader = (IMAGE_DOS_HEADER*)imageBase;
IMAGE_NT_HEADERS ntHeaders = dosHeader->e_lfanew;
IMAGE_OPTIONAL_HEADER optHeader = &ntHeaders->OptionalHeader;
void (*EntryPoint) (void) = (void (*EntryPoint) (void) ) optHeader->AddressOfEntryPoint + optHeader->ImageBase;

/* プログラムのエントリーポイントを
呼び出す */ EntryPoint();
```

同様に、他のヘッダーを解析して、セクション情報、デバッグ情報、シンボリック情報など、ファイルから必要なあらゆる種類の情報を抽出することができます。カーネルデバッガやユーザー モード デバッガは、通常、デバッグを容易にするためにシンボリック情報やデバッグ情報を使用します。言い換えれば、デバッグ情報を含む(または含まない)PEイメージを構築することができます。デバッグ情報を含むように構築した場合、リモートでデバッガを取り付けてソースレベルのデバッグを行うことができます。

実行ファイルの中に異なるセクションがあると、実行ファイルの解析や記述が容易になります。これらのセクションは、データを格納するための一貫した場所と方法を提供し、ヘッダからそれらを参照することができます。たとえば、PEファイルには、実際のリソース(文字列テーブル、ビットマップイメージ、プログラム情報、カーソルなど)を格納する.rsrcセクションがあります。リソースを見つけるためには、OptionalHeader->DirectoryEntries [IMAGE_DIRECTORY_ENTRY_RESOURCE] のディレクトリエントリを解析するだけで、.rsrcセクション内のリソースデータを指すリソースツリー構造への RVA(相対ポインター)が得られます。重要なのは、実行ファイルのフォーマットはPEファイルの仕様で定義された特定のフォーマットであるということです。フォーマットが決まっているので、ファイルから情報を得るための標準的な方法があるのであります。

GCCやCL (マイクロソフトのコンパイラ) のような多くのコンパイラは、プログラマがセクションを定義することができます。つまり、プログラマーは自分でセクション名を定義し、必要なものをそのセクションに入れることもできるのです。オペレーティングシステムのカーネルや幹部は、通常、異なる目的のために特別なセクションを定義します。例えば、LinuxもWindowsも、特別なセクションを定義しています。<.INITセクションには、1回限りの初期化コードとデータが格納されています。初期化が完了すると、OSカーネルはこのセクションを解放し、他の用途に再利用することができます。

共通部分

1. 異なるオブジェクトファイルやアーキテクチャに共通するセクションの名前とタイプがあります。これらのセクションを認識し、何に使用されているかを知ることは重要です。それらは以下の通りです。
 2. .text
 3. .data
 4. .bss
 5. .rodata

.textセクションは、プログラムコードを含むセクションに与えられる一般的なセクション名です。これはコードセグメントとも呼ばれています。システムによっては、このセクションに書き込みができないように、読み取り専用になっている場合がありますが、これはコードの自己修正を妨げることになります(これは通常、推奨されません)。

.dataセクションは、その名の通り、プログラムが使用するスタティックデータやグローバルデータを格納します。常に書き込み可能です。

.bssセクションは、.dataセクションの一部で、通常、ゼロに初期化された静的に割り当てられたデータに使用されます。このセクションは

.bssセクションは、OSローダーによって常にクリアされるため、その中のデータはすべてゼロになります。ウィキペディアによると、「.bss」という名前は、当初、United Aircraft Symbolic Assembly ProgramのBlock Started by Symbolを意味していました。.bssセクションにはすべてのスル変数が含まれているため、オブジェクトファイルの中でスペースを取りません。

.rodataセクションには、読み取り専用の静的に割り当てられたデータが格納されています。LinuxやUnixの環境でよく見られます。

一時データ用のセクションがないことに注意してください。一時変数はスタックに格納されるので、プログラムファイルに格納する必要がないことを思い出してください。

Microsoft Visual Cのカスタムセクション

1. Microsoftのコンパイラには、プログラマが特定のセクションやカスタムセクションにデータやコードを配置する際に使用できるプラグマディレクティブがいくつか用意されています。これらは

2. alloc_text
3. code_seg
4. const_seg
5. data_seg
6. bss_seg
7. init_seg
8. section

プログラムローダは、プログラムが持つ特別なセクションを気にする必要はなく、メモリにロードすることだけを考えればよい。プログラム（つまりプログラマー）が責任を負うのです。

ここでは、特別なセクションに関数を追加するためにalloc_textを使用する例を示します。

```
error_t DECL mmInitialize (SystemBoot* mb)
{
    return SUCCESS;
}
```

```
#pragma alloc_text (.init", mmInitialize);
```

上の例では、mmInitializeが.initセクションに追加されます。これは、一部のオペレーティング・システム・カーネルやエグゼクティブが使用している便利な戦術です。例えば、オペレーティング・システムのカーネルやエグゼクティブは、初期化コードやデータを、特別な.initセクションです。初期化が完了すると、OSはそのセクションを解放してメモリの一部を取り戻すことができます。

象徴的な情報

記号情報とは、プログラマーがアドレスの名前として与える記号のことです。例えば、次のような関数を呼び出す場合

printf()では、コンパイラやリンクはどのようにして何をすべきか知っているのでしょうか？もう少し詳しく見てみましょう。

"printf"は、ライブラリで定義されている関数のシンボルです。printf()」を呼び出すと、コンパイラがビルド時に管理しているシンボルテーブルにシンボル「printf」が追加されます。関数の名前がシンボルであることに注目してください。同様に、スタティック変数やグローバル変数もシンボルです。アドレスにつける名前（アセンブリ言語のラベルのようなもの）はすべてシンボルであることができます。このように、シンボルは2つのものを持っています。名前とアドレスです。

printfのコードを含むライブラリをリンクせずにビルドした場合、コンパイラはコード全体を機械語に翻訳することができないため、最終的な実行ファイルを出力することができません。つまり、アセンブリのように、以下のようなコードでは、関数が何であるかわからないと何もできないのです。

コール _printf

アセンブリは、シンボル _printf のアドレスを知らないと、この命令を完全にアセンブルできません。シンボルについて何も知らないと、アセンブリはアドレスを知ることができません。この問題を解決するために、シンボルが外部にあることを宣言し、アセンブリやコンパイラは実行ファイルではなくオブジェクトファイルを出力します。命令を機械語に部分的に変換しますが、次のような形になります。

0xe8 _printf

これにより、別のプログラム（リンカー）を使ってシンボルを解決することができます。リンカは、異なるオブジェクトファイルやライブラリのエクスポートシンボルテーブルを見て、シンボル「_printf」を探します。シンボルが見つかれば、リンカーはファンクションコードのアドレスを取得し、そのアドレスでマシンコードを更新して、最終的に実行可能なプログラムを適切にリンクして出力します。シンボルが見つからない場合は未解決となり、リンカーは有名な「未解決の外部シンボル」というエラーを出します。

実行イメージの記号情報は、デバッガが人間が読める情報（関数や変数）名を表示するために使用できますが、その代償としてプログラムのファイルサイズが大きくなります。

シンボルに関する追加情報をシンボル名自体に「格納」する方法はさまざまです。これは、ビルド環境や呼び出し規則によって異なります。標準的なC言語の呼び出し規則はCDECLで、すべての名前の前にアンダースコアを付けるだけです。例えば、"printf()"を呼び出した場合、そのCDECLシンボル名は"_printf"となります。C++のシンボリック名は、コンパイラによって異なり、名前以外にも多くの情報（戻り値のデータ型やオペランドの型、名前空間、クラス、テンプレート名など）を保持しています。例えば、CL（マイクロソフトのコンパイラ）の「void h(void)」という関数は、シンボリック名「?h@YAXXZ」に変換されます。ここでは、名前のマングリング形式の詳細については触れない。ここで面白いことに気がつきました。C言語のシンボリック名には、戻り値のデータ型やオペランドに関する情報は何も格納されていませんが、名前のマングリングによるC++のシンボリック名には格納されています。これは理解できますが、言語間の多くの違いの1つです。Cコンパイラでは、異なるオペランド型やオペランド数の関数をエラーなしで呼び出すことができますが（検出されると警告が出るかもしれません）、C++コンパイラではエラーが出ます。

テーブルのエクスポートとインポート

プログラムライブラリやオブジェクトファイル内のシンボルは、他のライブラリやプログラムで使用するためにエクスポートすることができます。エクスポートされたシンボルは、コンバイラやリンカーに、それぞれのシンボルをエクスポートテーブルに追加するように指示するだけです。プログラム・ファイルや共有ライブラリ（Windows DLL）は、他のプログラムやデバッガが使用するためにシンボルをエクスポートすることができます。同様に、プログラムファイルは使用するためにシンボルのインポートを要求することができます。ここで、上記のprintf()の例を完成させることができます。

Microsoft C Runtime Libraryは、プログラムファイルとともに読み込まれる共有ライブラリです。オペレーティングシステムやエグゼクティブは、プログラムのインポートテーブルを見ることで、プログラムファイルの動作に必要なDLLを知ることができます。デフォルトでは、CL(マイクロソフト社のコンパイラ)は、インポートテーブルを含むMicrosoft C Runtime Libraryのインポートスタティックライブラリとリンクするので、シンボルが追加され、それぞれのDLLがテーブルに含まれます。オペレーティングシステムまたはエグゼクティブは、プログラムが必要とするすべての共有ライブラリファイルをまだメモリにロードしていない場合は、プログラムファイルのインポートアドレステーブル(IAT)をこれらの他のDLLの関数のアドレスで更新する必要があります。ロードされたMicrosoft C Runtime Library DLLには、_printfのコードが含まれているだけでなく、シンボル_printfもエクスポートされているため、OSはランタイム中にこれらをリンクします。(この点については後ほど詳しく説明します)。

したがって、プログラムファイルから「printf()」を呼び出すと、ジャンプテーブルが呼び出され、更新されたIATアドレスが呼び出され、CランタイムライブラリDLLの「_printf」関数が呼び出されます。

これまでに、プロセス、スレッド、タスクについて説明し、プログラムファイルとは何か、どのように動作するのかを見てきました。本章の目標は、複数のプロセスやタスクをロード、実行、管理できるようになります。次はそれについて見てていきましょう。

Process Management

プロセスマネジメントとは、ソフトウェアシステムにおけるプロセスの管理です。先にプロセスを「メモリ上のプログラムまたはプログラムの一部」と定義しました。つまり、プロセスを管理するということは、メモリ上のプログラムの複数のインスタンスを協調して管理することを意味します。これは最近のOSでは一般的に要求されていることで、カーネルやエグゼキュティブに実装されています。プロセス管理をサポートするオペレーティングシステムは、マルチタスクオペレーティングシステムと考えられます。

表現

OS設計者は、プロセスを管理するために、OSの設計基準や必要なシステムリソースを考慮して、プロセスをどのように表現するのが最適かを判断する必要があります。プロセスは次のように構成されています。

メモリ上の実行ファイルのイメージ（マシンコードとデータ）。

プロセスが使用しているメモリとその仮想アドレス空間

プロセスを表現するための記述子

プロセスの状態情報（レジスタ、スタック、属性など）

オペレーティングシステムは、プロセスを管理し、システムリソースを要求したプロセスに公平に割り当てることが求められます。それぞれを詳しく見てみましょう。

メモリ上の実行ファイルのイメージ

実行可能なプログラムは、プログラムの読み込みや管理を容易にするために、ディスク上にファイルとして格納されている。プログラムをロードするには、オペレーティング・システム・ローダーがファイルをメモリにロードする。ローダーは、ファイルの種類（オペレーティングシステムが扱うことのできる実行可能ファイルでなければならない）を理解し、場合によってはこれらのファイルの種類の機能（リソースやデバッグ情報など）をサポートすることもできなければならない。

メモリ上の実行ファイルのイメージは、イメージのマシンコードとデータの現在の表現であり、ある時点でメモリ上にどのように表示されているかを示しています。ここで「イメージ」という言葉を使っているのは、メモリ上のものの「スナップショット」を表している。例えば、カメラで大きなバイトの配列を見て写真を撮るようなものです。バイトの配列は、マシンコードであったり、データであったり、そのどちらでもないものである。プログラムの命令だけが知っています。

プログラムイメージの中には、他のプログラムやOSにとって有用なデータが含まれている場合があります。例えば、プログラムファイルにはデバッグ情報が含まれていることがあります。例えば、プログラムファイルにはデバッグ情報が含まれており、プログラムにデバッガを取り付けて、その情報を利用することができます。

要するに、OSがファイルを実行するためには、ディスクからメモリのどこかにファイルをロードできる必要があります。これは、ファイルをそのままメモリにロードするようなものです。オペレーティングシステムや他のプログラムは、その後、プログラムファイルから必要な有用なデータを得ることができます。

プロセスが使用しているメモリとその仮想アドレス空間

プロセスは通常、オペレーティングシステムと同様に、動的にメモリを割り当て、スタックスペースを使用するコールを持っています。オペレーティングシステムは、プロセスが使用するプロセススタックヒープメモリのためのスペースを割り当てる必要があります。例えば、オペレーティング・システムは通常、すべてのプロセスにデフォルトのスタック・サイズを割り当てます。しかし、プロセスの実行ファイルは、プロセスが必要とする場合、より大きなスタックスペースを割り当てるようオペレーティングシステムに指示することができます。

プロセスのヒープは違います。スタックはプロセスを実行する前にオペレーティングシステムによって割り当てますが、ヒープはそうではありません。その代わり、各プロセスはユーザー モードで独自のヒープアロケータを持っています。これはCランタイム・ライブラリ（CRT）に実装されており、malloc、free、realloc、brk、sbrkの各関数の使い慣れたインターフェイスを使用しています。CRTとリンクしているプログラムは、これらの関数を呼び出してメモリを割り当てるすることができます。しかし、CRTとリンクされていないプログラムは、独自のヒープアロケータを実装するか、ヒープアロケータを実装している他のライブラリとリンクする必要があります。

CRT ランタイムは、ユーザモードのヒープアロケータ（通常はフリーリスト）を実装しています。C関数のmallocは、System APIを使用してOSを呼び出すbrkを呼び出します。C関数のbrkは、必要に応じてヒープを拡張するために、より多くの仮想メモリを割り当てるために、OSを呼び出します。

簡単に言うと、ユーザモードのヒープは次のように動作します。プログラムがmallocを呼び出し、そのmallocがbrkを呼び出し、brkがSystem APIを使ってOSを呼び出し、ヒープ用の仮想メモリを確保します。mallocおよびfree関数群は、独自のユーザモードヒープアロケータを実装しています。これらの関数は、仮想アドレス空間からメモリを割り当てたり解放したりするためにOSを呼び出すだけです。

プリエンプティブ・マルチタスクでは、すべてのプロセスが独自の仮想アドレス空間を持ちます。これは、すべてのプロセスが自分のページディレクトリと関連するページテーブルを持たなければならないことを意味します。プロセス固有の情報を管理するために、PCB (Process Control Block) を使用します。次にそれを見てみましょう。

プロセスを表現するための記述子

プロセスコントロールブロック（PCB）は、プロセスやタスクに関する情報を格納するために使用されるデータ構造です。PCBには、割込み記述子ポインタ、ページディレクトリベースレジスタ（PDBR）などの情報が含まれます。保護レベル、実行時間、プロセスの状態、プロセスフラグ、VM86フラグ、優先度、およびプロセスID（PID）。PCBはより多くの情報を含んでいるかもしれません、実際にはOSに依存しています。

オペレーティングシステムは、プロセスを管理するためにPCBのリンクされたリストを使用することができます。新しいプロセスを作成するとき、オペレーティングシステムは新しい仮想アドレス空間を割り当て、イメージをロードしてマッピングし、新しいPCB構造をリストに添付する必要があります。スケジューラは、実行するプロセスを決定し、現在の状態を保存するためにPCBリストを使用します。

プロセス状態情報

プロセス状態の情報には、ある時点でのプロセスの全レジスタ状態、インメモリ状態、入出力要求状態などが含まれます。プロセスの状態は、タスクを切り替える際にPCBに保存されます。これは、マルチタスクOSの心臓部であるスケジューラによって行われます。さらに、プロセスの現在の実行状態は、オペレーティング・システムによるプロセスの実行を制御するために使用されます。

最も単純なケースでは、状態はRUNNINGかNOT RUNNINGのどちらかです。このモデルでは、作成されたばかりのプロセスは、NOT RUNNINGキューに格納され、実行中のときだけRUNNINGと表示されます。NOT RUNNINGとなったプロセスは、RUNNINGプロセスが終了するか、スケジューラ内のプロセスディスパッチャに割り込まれるまで、メモリ上に存在しますが、待機状態となります。

3つの状態からなるプロセス管理モデルでは、プロセスは「実行中」、「準備中」、「ブロック中」のいずれかになります。RUNNINGプロセスが、プロセスの待ち時間を必要とするものへのアクセスを要求した場合(I/O要求など)、オペレーティングシステムはプロセスをRUNNINGからBLOCKEDに変更することができます。要求が実行できるようになると、プロセスはRUNNINGまたはREADYのいずれかの状態に移行することができます。READY状態のプロセスは、プロセス・ディスパッチャによる実行の準備ができていることを意味するだけです。RUNNING状態のプロセスは、すでに実行されています。

最後のモデルは、「5つの状態のプロセス管理モデル」です。このモデルでは、5つの状態を利用しています。SUSPEND BLOCKED」、「BLOCKED」、「SUSPEND READY」、「READY」、「SUSPENDED」です。

スケジューリング

スケジューラーとは、OSのカーネルやエグゼクティブのコンポーネントで、タスクの切り替えやCPU使用率の割り当てを行うものです。オペレーティングシステムでは、次に実行するタスクを決定するためにスケジューリングアルゴリズムを採用しています。一般的なスケジューリングアルゴリズムとしては、先入れ先出し、最短残り時間、固定優先プリエンプティブ、ラウンドロビン、マルチレベルキューなどがありますが、これらに限定されるものではありません。WindowsとLinuxの両方で使用されている最も一般的なアルゴリズムは、マルチレベルのフィードバックキューです。

Basic process management support

これで、基本的なプロセス管理サポートを実装することができます。目標はシンプルさなので、vm86タスクのサポートやI/Oリソースの割り当てなど、高度なマルチレベルのフィードバックシステムは実装せず、シンプルだが効率的なスケジューラに焦点を当てます。

1. そのためには、サポートを追加するために何をしなければならないのか、その目的を考えてみましょう。
2. Load and parse an executable image into memory;
3. Manage a list of PCBs for processes;
4. Support user mode tasks;
5. Support multiple virtual address spaces
6. Allocate stack space for each process; default size can be 4k;
7. Select a scheduling algorithm and implement task switching

これはマルチタスクに対応するための目標です。プロセスはユーザー モードのプロセスになります。しかし、マルチタスクはプロセス管理とスケジューリングの両方に依存します。このため、本章では、マルチタスクをサポートするフレームワークの構築に焦点を当てますが、1つのスレッドを持つ1つのプロセスのみを許可します。これは、次の章でスケジューラーを実装する際に拡張されます。

Process Control Block

私たちのシステムのPCB構造はシンプルなものになります。

```
#define PROCESS_STATE_SLEEP 0
#define PROCESS_STATE_ACTIVE 1
#define MAX_THREAD 5
typedef struct _process {
    int id;
    int priority;
    pdirectory* pageDirectory;
    int state;
} process;
/* typedef struct _process* next; */
/* thread* thread List; */
threads[MAX_THREAD];
/*
 * note: 以下のような情報を追加することができます。
 * - LDT記述子 [ 使用されている場合 ]
 */
```

```
- 使用されているプロセッサ数  
- ユーザータイムとカーネルタイム  
- 実行オプションなど  
*/  
}process;
```

この構造にさらに追加することができますが、実際に必要なのは上記のものだけです。この構造体には、プロセスID (PID)、優先度、仮想アドレス空間が格納されていることに注目してください。コメントされている2つのエントリは、完全性を保つためだけに用意されています。一般的なOSでは、これらはプロセスとスレッドのリンクリストになるはずです。しかし、これにはカーネルのヒープアロケータが必要ですが、我々は書いていません。簡単にするために、プロセス内の5つのスレッドオブジェクトを配列として保存します。

最後に必要なのは、スレッドを処理する方法です。すべてのプロセスは、エントリーポイントで実行を開始する最大1つのスレッドを持っています。

```
typedef struct _thread
{
    process* parent;
    void*     initialStack;
    void*     stackLimit;
    void*     kernelStack;
    uint32_t  priority;
    int       state;
    trapFrame frame;
}thread;
```

スレッド構造体は、プロセス内のスレッドに関する一般的な情報を格納します。この構造体には、親プロセスへのポインタ、スレッドスタック、優先度、状態（実行中かどうか）、およびトラップフレームに関する情報が格納されています。トラップフレームは、実行中のスレッドの現在のレジスタ状態を格納します。

```
typedef struct _trapFrame
{
    uint32_t esp;
    uint32_t ebp;
    uint32_t eip;
    uint32_t edi;
    uint32_t esi;
    uint32_t eax;
    uint32_t ebx;
    uint32_t ecx;
    uint32_t edx;
    uint32_t flags;
/*
    注：これにさらにレジスターを追加することができます。
    完全なトラップフレームのためには、追加する必要があります。
    - デバッグレジスター
    - セグメントレジスター
    - エラー状態 [もしあれば]
    - V 86 モードセグメントレジスタ [使用されている場合]
*/
}
```

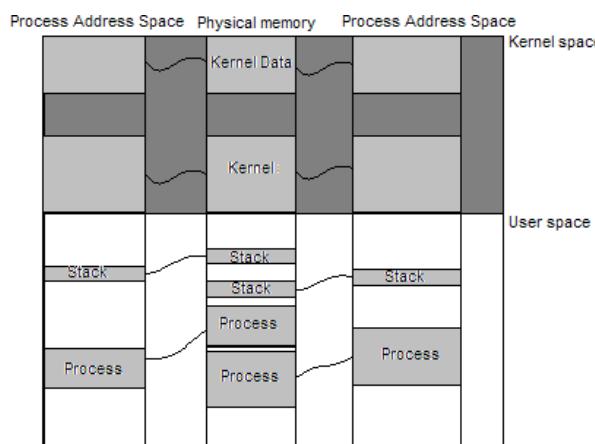
```
}trapFrame;
```

本章では、マルチタスクをまだ実装していないので、トラップフレーム構造はありません。しかし、次の章では、各スレッドの現在の状態を保存するスケジューラを開発するので、トラップ・フレーム構造をより多く使用することになります。

Virtual Address spaces

複雑なのは、複数の仮想アドレス空間をサポートしたい場合です。各プロセスのアドレス空間は、4GBのアドレス空間全体で構成されており、カーネルコードとデータは2GBに配置されています。プロセスを切り替える際には、アドレス空間を切り替える必要がありますが、アドレス空間の下位2GB（「ユーザーランド」）のみを切り替えることができます。言い換えば、ユーザー mode のプロセスが実行されているとします。タスクを切り替えるためには、カーネル内のスケジューラを何らかの方法で呼び出す必要があるのです。しかし、これはカーネルコードが同じアドレス空間になければならないことを意味しています。もしうでなければ、即座にクラッシュしてしまいます。

この問題を解決するためには、カーネルコードをすべてのプロセスのアドレス空間にマッピングする必要があります。どうやってそんなことができるのかと思われるかもしれません。しかし、複数の仮想アドレスがメモリ上の同じ物理フレームを参照できることを考えると、より明確になるかもしれません。言い換えば、カーネルのスタックとコードを両方のアドレス空間にマッピングすることができます。次の図をご覧ください。



上の画像では、2つの仮想アドレス空間と物理アドレス空間が表示されています。プロセススタックの位置とコードの位置が、物理メモリの異なる位置を共有していることに注目してください。つまり、仮想メモリマネージャを使って、同じ基本的な仮想アドレスの位置を、異なる物理アドレスフレームにマッピングしているのです。ここで、カーネルについて考えてみましょう。カーネルは、単一のアドレス空間を持つ環境で起動します。カーネルは、初期化プロセスにおいて、自分自身を自分のアドレス空間にマッピングします。問題を防ぐためには、カーネル空間を他のプロセスのアドレス空間にもマッピングできるようにしておく必要があります。カーネルはすでに自分のアドレス空間にマッピングされており、物理メモリのある場所に配置されています。つまり、カーネルは他のプロセスのアドレス空間にも自分自身を再マッピングすることができます。

できます。

シリーズのカーネルでは、カーネル自身が1MBの物理的な領域から3GBの仮想的な領域にマッピングされます。カーネルは、自分自身を各プロセスのアドレス空間にマッピングするために、すべてのプロセスの3GBの領域を1MBの物理的なものにマッピングしなければなりません。カーネルとカーネルスタックは、すべてのプロセスのアドレス空間の同じ場所にマッピングされなければなりません。

オペレーティングシステムでは、カーネル全体ではなく、カーネルの一部をプロセスアドレス空間にマッピングすることもできます。これは大規模なシステムではよくあることです。

アドレス空間の管理

1. 異なるアドレス空間の仮想ページをマッピングすることができるようになります。具体的には、次のようなことができるようになります。

2. Create a page table from any page directory
3. Map any physical address to virtual address from any page directory
4. Get the physical address of any virtual mapping from any page directory
5. Create new address spaces

シリーズの仮想メモリマネージャは、現在この機能をサポートしていません。しかし、すぐに実装することができますので、今からでも実装してみましょう。

ページテーブルの作成

ページテーブルを作成するために必要なことは、空きフレーム（ページテーブルは1024個のPTEで構成され、4096バイトが1ページのサイズであることを思い出してください）を割り当て、それをページディレクトリのPDEのフレームに追加することです。Virt >> 22では、仮想アドレスからディレクトリのインデックスを取得するだけです。pagedir [directory_index]のPDEが0であれば、このページテーブルは存在しないことがわかるので、物理メモリマネージャを使って割り当てます。存在していれば、割り当ての必要はありません。最後に、ページテーブルをクリアして、現在ビットを0（存在しない）にします。

```
int vmmngr_createPageTable (pdirectory* dir, uint32_t virt, uint32_t flags) { pd_entry* pagedir =
    dir->m_entries;
    if (pagedir [virt >> 22] == 0) { void* block =
        pmmngr_alloc_block(); if (!block)
            return 0; /* デバッガーを呼ぶべき */

        pagedir [virt >> 22] = ((uint32_t) block) | flags; memset ((uint32_t*)
            pagedir[virt >> 22], 0, 4096);
        /* ページテーブルをディレクトリにマッピングする。
        vmmngr_mapPhysicalAddress (dir, (uint32_t) block, (uint32_t) block, flags);
    }
    return 1; /* 成功 */
}
```

この機能により、任意のページディレクトリのページテーブルを作成することができます。

物理アドレスのマッピング

次に不足している機能は、異なるページディレクトリの物理アドレスと仮想アドレスをマッピングできるようにすることです。これは簡単です。

```
void mapPhysicalAddress (pdirectory* dir, uint32_t virt, uint32_t phys, uint32_t flags) {
    pd_entry* pagedir = dir->m_entries;
    if (pagedir [virt >> 22] == 0)
        createPageTable (dir, virt, flags);
    ((uint32_t*) (pagedir[virt >> 22] & ~0xffff))[virt << 10 >> 10 >> 12] = phys | flags;
}
```

この関数は、以前に仮想メモリマネージャに実装した基本的な機能を踏襲しています。有効なページテーブルがあるかどうかをテストし、存在しないとマークされていればページテーブルを作成します。最後の行ではマッピングを行います。

この機能により、任意の仮想アドレス空間の物理アドレスと仮想アドレスをマッピングすることができます。

物理的アドレスの取得

次に不足している機能は、先ほどとは逆に、特定のアドレス空間から任意の仮想アドレスの物理アドレスを取得することです。

```
void* getPhysicalAddress (pdirectory* dir, uint32_t virt) {
```

```
    pd_entry* pagedir = dir->m_entries;
    if (pagedir [virt >> 22] == 0)
        return 0;
    return (void*) ((uint32_t*) (pagedir[virt >> 22] & ~0xffff))[virt << 10 >> 10 >> 12];
}
```

この関数は、その仮想アドレスに有効なページテーブルがあるかどうか（存在するかどうか）をテストし、PDEとPTEを参照解除して物理フレームを返します。

新しいアドレス空間の作成

各プロセスは、それぞれの仮想アドレス空間で動作します。そのためには、複数のアドレス空間を作れるようにする必要があります。

```
pdirectory* createAddressSpace ()  
{ pdirectory* dir = 0;  
  
/* ページディレクトリの確保 */  
dir = (pdirectory*) pmmngr_alloc_block ()です。  
if (!dir)
```

```

0を返す。

/* メモリをクリアする（すべてのページテーブルを存在しないものとしてマークする） */ memset (dir, 0, sizeof (pdirectory) ) 。
dirを返す。
}

```

この関数のシンプルさに注目してください。この関数が行なうことは、ブロックを割り当ててクリアするだけです。これは、ページディレクトリがアドレス空間を表し、1つのページディレクトリが4096バイトであることから理解できます。クリアすることで、すべてのPDEで現在のビットを0にしていることになります。

プロセスを実行する際には、作成したばかりの新しいアドレス空間に切り替えることができなければなりません。つまり、この新しいページディレクトリをPDBRにロードする必要があるのです。この機能はPMMですでに実装されています。しかし、空のページディレクトリをそのままPDBRにロードすると、直後に必ずトリプルフォールトが発生してしまいます。この原因は簡単で、カーネルコードやスタックがこの新しいアドレス空間にマッピングされていないからです。

これを解決するには、カーネル空間をマッピングする必要があります。興味深いことに、次のようにすれば、現在のページディレクトリ（PDBRに格納されている）をこの新しいアドレス空間にコピーすることができます。

```
memcpy (dst->m_entries, cur->m_entries, sizeof(uint32_t)*1024);
```

必要なのはこれだけです。ページテーブルはすでに元のページディレクトリにマッピングされているので、ページテーブルのコピーを気にする必要はありません。以上の作業により、効果的にアドレス空間のコピーが作成され、カーネルのページテーブルが両方のアドレス空間にマッピングされ、これが必要な状態になります。

Creating a thread

スレッドを作成するためには、まず `createThread` 関数に必要なものと、スレッド作成が実際に意味するものを決定する必要があります。スレッドとは単一の実行経路であると定義したこと思い出してください。これを踏まえて、必要なのはエントリーポイント関数です。この関数が完了すると、オペレーティングシステムを呼び出してスレッドを終了させます。これは通常、スレッドの作成と終了を単純化するために、システムAPI（Win32 APIなど）によって行われます。

スレッドを作成するために必要なのは、スレッド構造体を割り当ててプロセスに追加することだけです。当初、この機能はこのデモのために実装する予定でしたが、25章のマルチスレッドに回すことになりました。

Process creation

プロセスを作成するためには、OSの専用ローダーコンポーネントがすでに存在している必要があります。ローダーコンポーネントは、実行ファイルの読み込みと解析、BSSセクションのクリア、セクションのアライメント、その他、ダイナミックリンクライブラリのダイナミックローディングなど、必要と思われることを行います。ローダーの作成は、特にPEのような複雑なファイルフォーマットの場合、複雑な作業となります。そのため、この章の目的であるプロセス管理に集中できるよう、このシリーズではよりシンプルなソリューションを選ぶことにしました。

1. プロセスをクレートするためには、プロセスとは何か、スレッドのそれとどう違うのかを明確に理解する必要があります。具体的に言うと
 2. Threads each have their own dedicated stack; the process itself does not have one.
 3. Each process must have at least one thread. This starts at the entry point of the process.
 4. Each process must have their own virtual address space. Threads in a process share the same address space as the process.
 5. Each process must be loaded from disk as an executable image. This is typically done using a separate loader component.
1. シリーズには専用のイメージローダーコンポーネントがないため、簡単にするために、`createProcess`という1つの関数でこれらのステップをすべて実行します。この関数は次のようなステップを踏みます。
 2. Load the executable file.
 3. Create the address space for the process.
 4. Create a Process Control Block (PCB).
 5. Create the main thread.
 6. Map the image into the process virtual address space.

これは1つの機能としてはかなりの量です。後で実行ファイルのロードを専用のローダに移すことができます。簡単にするために、このルーチンはブートローダと同じ基準を想定しています。つまり、ロードされるイメージは、セクタアラインされたセクション（/align:512フラグを使用）を持っていなければならず、Microsoft Windowsのランタイムライブラリとリンクされていてはならないのです。将来的にはこの機能をデモに追加するかもしれません、これはローディングコードを複雑にしますし、先に述べたように、通常はローダコンポーネントが処理します。

プロセス・アドレス・スペース構造

現在、シリーズOSのカーネルは、アイデンティティマップドメモリーの1MB以下に多くのカーネル構造がロードされています。これには、カーネルスタック、初期ページディレクトリテーブル、ページテーブルなどが含まれます。また、DMAC（Direct Memory Access

Controller) のメモリ領域もこの領域に配置されている可能性があります。また、カーネルが他のメモリ領域（ディスプレイメモリなど）も利用していることも考慮しなければなりませんが、それらもこのアイデンティティマップド領域にあります。

これはすべて、単純化のために行われたものです。一般的なカーネルは、PIC (Position Independent Code) を使って、カーネルメモリ内のカーネルスタックと初期ページテーブルを初期化します。PICは、上位のカーネルが他の場所でロードされたときに起動するためのものでもあります。

物理的なベースアドレス。これを正しく行うのは難しいので、今回のシリーズでは採用しませんでした。しかし、その結果、1MB以下のカーネル構造がいくつかできてしまい、混乱してしまいました。

いろいろなものを移動させるよりも、0~4MBをカーネルモード専用に確保するのがベストな選択だと思いました。これにより、カーネルは何も手を加えずに機能し続けることができ、ディスプレイ出力やその他の基本的なことのためにメモリをリマップすることも問題ありません。つまり、アドレス空間は次のようにになります。

```
0x00000000-0x00400000 - カーネルの予約
0x00400000-0x80000000 - ユーザーランド
0x80000000-0xfffffff - カーネルの予約
```

つまり、すべてのプロセスは、4MBと2GBの領域内にイメージベースを持たなければなりません。ここでは、すべてのユーザー モード プロセスのベースアドレスとして4MBを使用します。最初の4MBはカーネルモードのページとして同一にマッピングされたままで（これはすでに行われています）、カーネル自体は3GBにマッピングされたままで。つまり、プロセスのすべてのページは、4MBと2GBの間のユーザー モード ページとしてマッピングされます。

プロセスの作成

それでは早速、関数を見てみましょう。これは、一般的にローダーで行われるソフトウェアを含んでいるため、かなり長いルーチンになっています。今回のデモでは、新しいアドレス空間を作るのではなく、現在のアドレス空間にイメージをロードしてマッピングしていますが、両方も実装されています。これは、このソフトウェアが一度に1つのプロセスしか実行できないように設計されているためです。第25章のマルチタスクでは、この点を変更します。

vmmngr_createAddressSpaceとmapKernelSpaceという2つの新しい関数に注目してください。これらは現在デモでは使用されていません。
vmmngr_createAddressSpaceとmapKernelSpaceという2つの新しい関数に注目してください。つまり、カーネルメモリ、スタック、ページディレクトリ、ディスプレイメモリを新しいアドレス空間にマッピングします。これらの機能は使用されていませんが、次の章で使用されます。

validateImage関数は、画像のヘッダーを解析し、サポートされているかどうかを確認するだけです。最後に、初期のスレッド構造を作成しますが、マルチスレッドには対応していません。プロセスごとに1つのスレッドを想定しており、1つのプロセスの1つのスレッドしか実行できません。

```
int createProcess (char* appname) {
```

```
    IMAGE_DOS_HEADER* dosHeader;
    IMAGE_NT_HEADERS* ntHeaders;
    FILE file;
    pdirectory* addressSpace;
    process* proc;
    thread* mainThread;
    unsigned char* memory;
    uint32_t i;
    unsigned char buf[512];
    /* ファイルを開く */
    file = volOpenFile (appname); if
    (file.flags == FS_INVALID)
    0を返す。
        if ((file.flags & FS_DIRECTORY) ==
            FS_DIRECTORY) return 0;

    /* 512バイトをバッファに読み込む */
    volReadFile ( &file, buf, 512);
    if (! validateImage (buf))
    { volCloseFile ( &file ); return
    0;
    }
    dosHeader = (IMAGE_DOS_HEADER*)buf;
    ntHeaders = (IMAGE_NT_HEADERS*)(dosHeader->e_lfanew + (uint32_t)buf);
    /* プロセスの仮想アドレス空間の取得 */
    //addressSpace = vmmngr_createAddressSpace (); addressSpace =
    vmmngr_get_directory ()です。
    if (!addressSpace)
    { volCloseFile ( &file );
    return 0;
    }
    /*
    カーネル空間をプロセスのアドレス空間にマッピングし
    ます。新しいアドレス空間を作る場合のみ必要
    */
    //mapKernelSpace (アドレススペース) です。
    /* PCBを作成する。
    proc = getCurrentProcess ();
    proc->id = 1;
    proc->pageDirectory = addressSpace;
    proc->priority = 1;
    proc->state = PROCESS_STATE_ACTIVE;
    proc->threadCount = 1;
    /* スレッド記述子の作成 */
    mainThread = &proc->threads[0];
    mainThread->kernelStack = 0;
    mainThread->parent = proc;
    mainThread->priority = 1;
    mainThread->state = PROCESS_STATE_ACTIVE;
    mainThread->initialStack = 0;
    mainThread->stackLimit = (void*) ((uint32_t) mainThread->initialStack + 4096);
    mainThread->imageBase = ntHeaders->OptionalHeader.ImageBase;
    mainThread->imageSize = ntHeaders->OptionalHeader.SizeOfImage;
    memset (&mainThread->frame, 0, sizeof (trapFrame));
    mainThread->frame.eip = ntHeaders->OptionalHeader.AddressOfEntryPoint
```

```
+ ntHeaders->OptionalHeader.ImageBase;  
mainThread->frame.flags = 0x200;  
/* 上で読み込んだ512ブロックと残りの4kブロックをコピーします。
```

```

memory = (unsigned char*)pmmngr_alloc_block();
memset (memory, 0, 4096);

memcpy (memory, buf, 512)。
/* 画像のメモリへのロード */
for (i=1; i <= mainThread->imageSize/512; i++) { if
(file.eof == 1)
    ブレークします。
    volReadFile (&file, memory+512*i, 512);
}

/* ページをアドレス空間にマッピングする */
vmmngr_mapPhysicalAddress (proc->pageDirectory,
    ntHeaders->OptionalHeader.ImageBase,
    (uint32_t) memory,
    i86_pte_present|i86_pte_writable|i86_pte_user)になります。

/* 残りの画像のロードとマッピング */
i = 1;
while (file.ef != 1) {
    新しいフレームを割り当てる /*。
    符号化されていない char* cur = (符号化されていない char*)pmmngr_alloc_block()。
/* ブロックを読む */ int
    curBlock = 0;
    for (curBlock = 0; curBlock < 8; curBlock++) { if
(file.eof == 1)
    ブレークします。
    volReadFile (&file, cur+512*curBlock, 512);
}
/* ページをプロセスのアドレス空間にマッピングする */
vmmngr_mapPhysicalAddress (proc->pageDirectory,
    ntHeaders->OptionalHeader.ImageBase + i*4096,
    (uint32_t) cur,
    I86_PTE_PRESENT|I86_PTE_WRITABLE|I86_PTE
    _USER);
i++;
}

/* ユーザ空間スタックの作成 ( プロセス esp = 0x100000 ) */
void* stack =
(void*) (ntHeaders->OptionalHeader.ImageBase)

+ ntHeaders->OptionalHeader.SizeOfImage + PAGE_SIZE) ; void*
stackPhys = (void*) pmmngr_alloc_block ();
/* ユーザプロセスのスタック空間をマッピングする。

vmmngr_mapPhysicalAddress (addressSpace, (uint32_t) stack, (uint32_t) stackPhys,
I86_PTE_PRESENT|I86_PTE_WRITABLE|I86_PTE_USER)になります。

/* 最終的な初期化 */ mainThread-
>initialStack = stack;
mainThread->frame.esp      = (uint32_t)mainThread->initialStack;
mainThread->frame.ebp      = mainThread->frame.esp;

/* ファイルを閉じてプロセスIDを返す */ volCloseFile(&file);
return proc->id;
}

```

Process execution

プロセスを実行するには、そのプロセスのメインスレッドからEIPとESPを取得し、ユーザー mode に落として実行すればよいのです。しかし、ここで問題が発生します。どのプロセスを実行すればいいのかをどうやって判断するのか？スケジューラがまだないため、一度に1つのプロセスしか実行できません。これは、現在どのプロセスで作業しているかを保存するグローバルプロセスオブジェクトを使用して行われます。GetCurrentProcess()は、このオブジェクトへのポインタを返します。そのメインスレッドからESPとEIPを取得し、プロセスのアドレス空間に切り替え、ユーザー mode に落として実行します。

enter_usermodeを呼んでいないことに注意してください。これは、ユーザー mode のソフトウェアがカーネルオンリーのページにアクセスできないためです。これを呼び出すと、ページフォルトになってしまいます。代わりに、ユーザー mode に落ちて、IRETDを使って直接プログラムを実行します。

```

void executeProcess ()
{ process* proc = 0; int
entryPoint = 0;
unsigned int procStack = 0;

/* 実行中のプロセスを取得 */
proc = getCurrentProcess();
if (proc-
>id==PROC_INVALID_ID)
return;
if (!proc->pageDirectory)
    を返すことができます。

/* メインスレッドのespとeipを取得 */ entryPoint =
proc->threads[0].frame.eip; procStack = proc-
>threads[0].frame.esp;
/* プロセスアドレス空間への切り替え */
asm cli
pmmngr_load_PDBR ((physical_addr)proc->pageDirectory).

```

```
/* ユーザモードでプロセスを実行 */
asm {
    mov    ax, 0x23          ; user mode data selector is 0x20 (GDT entry 3). Also sets RPL to 3
    mov    ds, ax
    mov    es, ax
    mov    fs, ax
    mov    gs, ax
    ;
```

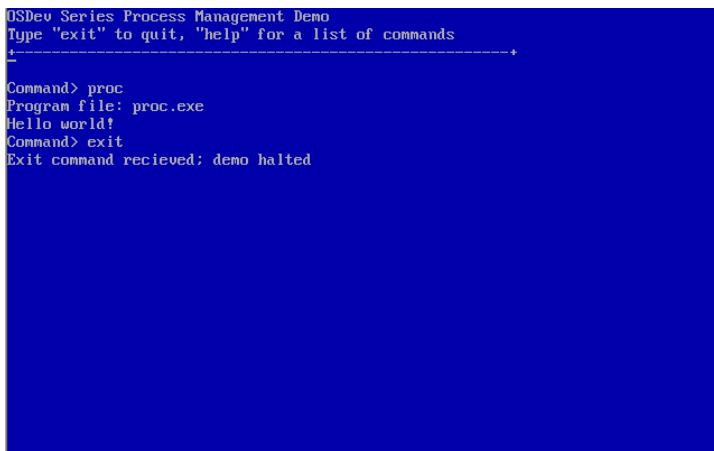
```

        ; スタックフレームの作成
        ;
        push    0x23          ; SS, notice it uses same selector as above
        push    [procStack]    ; stack
        push    0x200         ; EFLAGS
        push    0x1b          ; CS, user mode code selector is 0x18. With RPL 3 this is 0x1b
        push    [entryPoint]   ; EIP
        iretd
    }

}

```

Demo



システムAPIを使用するusermodeプロセスの実行

概要

これは、開発中のオペレーティングシステムにとって重要なマイルストーンです。このマイルストーンは、インターラクティブ性とセルフホスティングシステムデザインの始まりを意味します。このデモでは、本章、メモリ管理の章、PEのロードの章で説明した内容を使用して、実行イメージ (Portable Executable format) をロードするproc (プロセス) コマンドを実装し、ユーザーランドの独自のアドレス空間にマッピングし、2つのシステムコールを通じてシステムAPIを使用してカーネルと対話します。ユーザーランドのプロセスがカーネルのテキスト端末を使って文字列を表示するために使用できるDebugPrintfと、プロセス自体を終了させるために使用されるTerminateProcessです。このシステムコールは、これまでの章で説明したソフトウェア割り込みを利用して実装されています。

Project "proc"

使用したusermodeプロセスはprocと呼ばれています。これは32ビットのPE実行可能イメージとして構築され、イメージベースは4MB、512バイトのセクションアラインメントです。参考までにこのプロジェクトのソースを紹介します。

void processEntry () { ... }

```

char* str="Hello world!"。
__asm {
    /* カーネル端末にメッセージを表示 */ mov ebx, str
    mov eax, 0
    int 0x80

    /* 終了 */ mov eax,
    1
    int 0x80
}
のために ( ; ; )。
}

```

このプロセスでは、メッセージの表示と終了にシステムコールが使用されていることに注目してください。これらのシステムコールは、これまでの章で実装したシステムAPIに追加されたものです。Int 0x80関数0はDebugPrintfを、Int 0x80関数1はTerminateProcessという新しい関数を呼び出しています。デモの機能を向上させるために、同様の方法でファイルの読み込みや入力などのシステムサービスを追加することができます。

TerminateProcessは、プロセスリソースをクリーンアップし、実行をカーネルコマンドシェルに戻す役割を果たします。int 0x80が実行されると、CPUはカーネルモードに移行し、CS、SS、ESPをTSSのそれぞれの値に戻すことを思い出してください。このように、システムコールが実行されるたびに、CPUはカーネルランドで同じアドレス空間を実行しています。これにより、TerminateProcessから直接カーネル関数を呼び出したり、カーネルコマンドシェルを呼び出したりすることができます。

```

extern [C] {
void TerminateProcess () {
    プロセス* cur = &_proc;
    if (cur-
        >id==PROC_INVALID_I
        D) return;

    /* スレッドの解放 */ int i=0;
    thread* pThread = &cur->threads[i];
}
}

```

```
/* スタックの物理アドレスを取得する。  
void* stackFrame = vmmngr_getPhysicalAddress (cur->pageDirectory,
```

```

        (uint32_t) pThread->initialStack);
/* スタックメモリをアンマップして開放する。
vmmngr_unmapPhysicalAddress (cur->pageDirectory, (uint32_t) pThread->initialStack);
pmmngr_free_block (stackFrame);
イメージメモリをアンマップして開放する */
for (uint32_t page = 0; page < pThread->imageSize/PAGE_SIZE; page++)
{ uint32_t phys = 0;
    uint32_t virt = 0;
    /* ページの仮想アドレスを取得 */
    virt = pThread->imageBase + (page * PAGE_SIZE);
    /* ページの物理的アドレスを取得 */
    phys = (uint32_t) vmmngr_getPhysicalAddress (cur->pageDirectory, virt) です。
    /* unmap and release page */ vmmngr_unmapPhysicalAddress
    (cur->pageDirectory, virt); pmmngr_free_block ((void*)phys);
}
/* カーネルセレクタの復元 */
__asm {
    cli
    mov eax, 0x10
    mov ds, ax
    mov es, ax
    mov fs, ax
    gs, ax sti
}

/* カーネルのコマンドシェルに戻る */ run ();
DebugPrintf ("z^w^~"); for (;;) ;
} // extern "C"

```

Bug report

過去のいくつかのデモでは修正されていますが、他のデモではまだ存在する可能性があるバグが再発しています。今後、このバグが存在する可能性のあるすべてのデモについて、修正プログラムをアップロードする予定です。このバグは vmmngr_initialize にあり、一部のデモでは物理メモリマネージャを初期化する前にこの関数を呼び出しており、これらのデモではカーネル空間のマッピングが不適切なため、ページフォルトやトリプルフォルトが発生する可能性があります。本章のデモでは（再び）解決されていますので、更新されたコードを main.cpp と mmngr_virt.cpp で確認してください。

Updated file list

sysapi.h - _syscallsが更新され、DebugPrintfとTerminateProcessが含まれるようになりました。

task.h - New.

task.cpp - 新規。

main.cpp - procコマンドを追加しました。また、VMMのバグフィックス。

mmngr_virt.h - 新しいアドレス空間関数です。

mmngr_virt.cpp - 新しいアドレス空間関数。また、VMMのバグフィックス。

image.h - PEのイメージ構造と定義。

proc/main.cpp - 新規です。

Conclusion

この章では、プロセス、スレッド、プロセス管理、および基本的なプロセス管理サポートの構築について説明しました。この章では、ユーザー モードのプログラムをディスクから実行するために必要なことをすべて網羅しており、これはオペレーティングシステムにとって大きな節目となりました。

次章では、本章で実装したプロセス管理機能をベースに、スケジューラーを構築し、プリエンプティブなマルチタスクのサポートを完成させます。

次の機会まで。

～Mike () です。

OS開発シリーズ編集部

Resources

以下のリンクは、より直接的で正確な情報を提供するために参照されたものです。さらなる情報を得るために参照してください。

[\[リンク\]](#)

[http://en.wikipedia.org/wiki/Scheduling_\(computing\)#Scheduling_disciplines](http://en.wikipedia.org/wiki/Scheduling_(computing)#Scheduling_disciplines)

[http://en.wikipedia.org/wiki/Process_management_\(computing\)](http://en.wikipedia.org/wiki/Process_management_(computing))

Additional links

以下のリンクは、このトピックに関連する追加のチュートリアルやリソースです。これらは、教材の補足として、あるいは異なるデザインを提供する際に役立つかかもしれません。もし、追加した方が良いと思われるリンクをご存じでしたら、ぜひ教えてください。また、リンク先には、次の章まで詳しく見ないマルチタスクの概念が含まれている場合があります。

[View](#)



第23章



[Index](#)



オペレーティングシステム開発シリーズ

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

"Controlling complexity is the essence of computer programming" - Brian W. Kernighan

25: Process Management 2

by Mike, 2015

1. Introduction

Welcome!

1. 前章では、プロセス間通信（IPC）、保護、リソースの割り当て、プロセス制御ブロック（PCB）、プロセスの実行状態、プロセスのアドレス空間など、プロセス管理の基本的なトピックについて詳しく説明しました。また、シングルタスクのサポートと基本的なシングルタスクの実装についても詳しく説明しました。本章では、前章の続きとして、マルチタスク、スケジューリング、セキュリティ、相互排除に重点を置いて、それぞれのトピックについてさらに詳しく説明します。具体的には、以下の項目について説明します。

2. Multithreading;
3. Multitasking;
4. Init and Idle Process;
5. Kernel/User Shared Data Space;
6. Mutual exclusion and Semaphores;
7. Introduction to Concurrent programming;
8. Scheduling algorithms;
9. Introducing to the MP Standard.

本章では、前章を読んでいただいたことを前提に、実際の設計や実装に焦点を当てた、より高度な内容をお届けします。前章と同様に、まずこれらのトピックの背後にある理論に触れ、次にユーザーランドプロセスに完全なマルチスレッドを実装するデモを紹介します。なお、本章ではMP規格について簡単に紹介していますが、詳細については後述します。MPサポートを実装するには、APICの適切なサポートが必要ですが、これは高度なトピックです。

2. Process State Management

本連載では、プロセスについて多くのことを説明してきましたので、今回は、プロセスの状態とプロセスの作成についてのおさらいです。前章では、プロセスを作成するための関数を実装しました。今回のデモでは、この関数を修正して、プロセスが適切に実行されるように、プロセス用の新しいタスクを作成します。プロセスのスケジューリングと密接な関係があるので、状態管理についても確認しておきましょう。

プロセスの状態とは、そのプロセスが採用している現在のアクティビティのことです。最低でも、プロセスは作成、実行、実行準備、終了が可能です。これだけでも4つの状態が考えられます。

新しい。プロセスの作成中です。

実行中です。プロセスが実行されています。

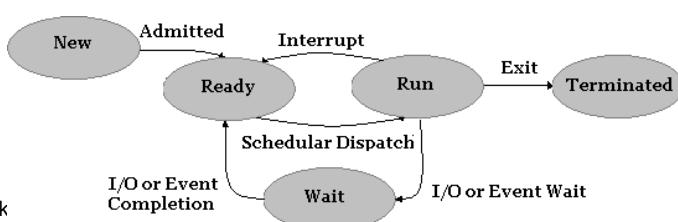
準備完了です。プロセスを実行する準備ができています。

終了しました。プロセスが終了しました。

これは良いスタートだと思います。しかし、もっと良い方法があります。例えば、あるプロセスが動作していて、そのプロセスがディスクから大きなファイルを読み取るリクエストを送ってきたとします。しかし、複数のプロセスが存在するシステムでは、ディスクはそのプロセスからの要求を処理するのに忙しくなる可能性があります。私たちのプロセスは、入出力要求が完了するまで待つ必要があります。別の例として、2つのプロセスがありますが、それらのプロセスはシグナルで相互に通信しているとします。プロセスは、シグナルが発生するのを待つ必要があります。これが5番目の状態になります。

待ちます。プロセスは、I/Oリクエスト、例外、またはシグナルの完了を待っています。

これらをまとめると、プロセスは次のような状態を経ることになります。



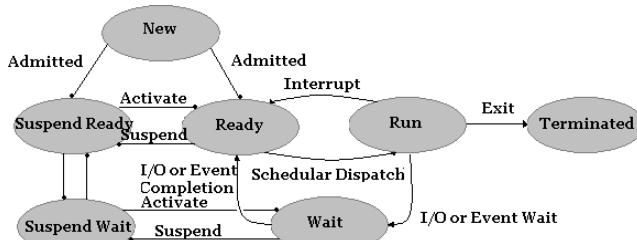
上の図は、現在の状態モデルを示しています。新しいプロセスは、システムのReadyキューに入れられます。が終了すると

スケジューラディスパッチャが実行するプロセスを選択すると、プロセスは実行状態になります。ここから、プロセスはいくつかの状態変化を取ることができます。割り込みや例外が発生した場合、スケジューラディスパッチャは他のプロセスに切り替える必要があり、その際にはプロセスをレディキューに戻すことになります。代わりに、プロセスがファイルから読み取ろうとした場合、プロセスはI/O要求を開始し、要求が完了するまでWaitキューに入れられます。I/O要求が満たされたら、プロセスはReadyキューに戻され、再びSchedulerディスパッチャによって選択されることになります。最後に、プロセスの実行中にいつでも、プロセスは終了します。

時には、プロセスをサスペンドすることが有効な場合があります。これは、プロセスをメモリから取り出し、その状態をディスクに保存するものです。これは、システムリソースを解放し、より優先度の高い他のプロセスを実行できるようにする場合に特に有効です。このためには、最低でもあと2つの状態が必要です。

サスペンドレディ。
サスペンド・ウェイ
ト。

これらを先ほどの図に加えると、以下のようになります。



「Ready」または「Wait」状態のプロセスは、システムのリソース需要に応じて一時停止することができます。デザインの必要性に応じて、この他にも様々な状態を追加することができますが、ほとんどの汎用OSでは、上記の状態図で十分です。

ここでは、Ready、Run、Terminatedの各状態についてのみ説明します。しかし、次のような実装も可能です。
付属のデモでスリープ機能を適切にサポートするために、Wait状態にしています。

3. Concurrent Programming

そこで、プロセスの状態と状態管理、プロセスの作成について見てみました。また、前章ではメモリ上のプロセスについて深く掘り下げました。最後に、マルチタスクについて説明します。マルチタスクの中心となるのは、次の章で紹介するスケジューラディスパッチャです。スケジューラディスパッチャは、プロセスの状態を移動させたり、プロセスの実行をスケジューリングする役割を担っています。このセクションでは、これらの機能を使用します。スケジューラの話をする前に、複数のスレッドが実行されているときのマルチタスクについて詳しく見ておきましょう。2つのスレッドやプロセスが同時に実行され、お互いにデータを共有する場合、2つの実行スレッド間のアクティビティを同期させることが重要になります。

同時進行とは、プロセスの現在の状態がわからないことを意味します。複数のプロセスが並行して実行され、互いにデータを共有している状態を「**同時実行**」といいます。同時実行プログラミングは、同時実行中のプロセスまたはスレッド間で共有リソースへのアクセスを同期させるために使用される一連の技術を定義します。

クリティカルセクションの問題

シングルコアのシステムでは、OSは各プロセスに少量の実行時間を割り当てます。システムは、同時に実行されているさまざまなプロセスを迅速に切り替えます。プロセスはいつでも中断することができます。また、並列実行をサポートするシステムでは、異なるプロセスの命令を同時に実行することができます。

現在のプログラミングの問題を見るために、次のような命令を持つ2つのプロセスを考えてみましょう。

Process A	Process B
mov eax, [count]	mov ebx, [count]
inc eax	dec ebx
mov [count], eax	mov [count], ebx

これらのプロセスを同時に実行する場合、スケジューラーが2つのプロセスを切り替える際に、何らかの順序でインターリーブされることになります。プロセスのインターリーブには様々な方法がありますが、1つの方法は次のようなものです。

```
mov eax, [count] inc
eax
mov ebx, [count]
dec ebx
```

```
mov [count], eax
mov [count], ebx
```

countが2つの異なるプロセス間で共有されている場合、ここで大きな問題に気づくかもしれません。実行順序が制御されていないため、スケジューラが2つのプロセスを切り替えることを決定するタイミングによって異なる結果が得られる可能性があり、countの値が有効であることを保証できません。どちらが先に変数を読み書きするかによって結果が変わります。これはレースコンディションと呼ばれています。

競合状態に対抗するには、他のプロセスが使用している間、変数を保護する必要があります。何らかの方法で2つのプロセスを同期させる必要があります。これは、クリティカルセクション問題の一部です。

複数のプロセッサを搭載したシステムでは、1つのプロセスを実行する際に、現在の実行状態と現在の命令ストリームがインターリープされるため、この問題はさらに悪化します。

問題。

並行して実行されるプロセスの同期を制御する方法が必要です。クリティカルセクションが要求された場合、クリティカルセクションが完了するまで、1つのプロセッサだけがクリティカルセクション内のコードを実行するようにしなければなりません。さらに言えば、クリティカルセクションに入っている間、他のプロセスやスレッドが実行されないようにしなければなりません。

基準は

Mutual Exclusionの略。あるプロセスがクリティカル・セクションで実行されているとき、他のプロセスはクリティカル・セクションで実行されていない。

プログレス。プロセスは、クリティカルセクションに入るまで、いつまでも待っているわけではありません。

バウンデッド・ウェイティング。そのクリティカルセクションに入るためのリクエストをしてから、実際にに入るまでの時間は拘束されていなければなりません。

セマフォ

では、どのようにして相互排除を実施するか。2つのプロセスの間に何らかの協力関係が必要です。プロセスAが共有リソース上で動作していて、プロセスBがそのリソースへのアクセスを必要とする場合、プロセスBには待ってもらいたい。しかし、プロセスAがリソースを使い終わったら、プロセスBがそのリソースを使えるようになったことを知らせる必要があります。このように、常に1つのプロセスだけが共有リソースを使用することができます。これが相互排除です。

私たちができることは、リソースが現在使用されているかどうかを追跡するために、別の変数を導入することです。この変数をロックと呼びます。そして、このロックを使って、他のリソースを追跡することができます。

ロックが1の場合、そのリソースは他のプロセスによって使用されています。ロックが0の場合、そのリソースは自由に使用できます。

この種のロックには特別な名前があります。それは「ミューテックス」と呼ばれるものです。ミューテックスは2つの値しか持たず、バイナリセマフォとも呼ばれます。必要なことを思い出してください。一方のプロセスが待機し、もう一方のプロセスがシグナルを送る必要があります。これらは、この章を通して使用する基本的な関数です。

```
atomic Wait (Semaphore S)
{ while (S <= 0)
    S.Queueにプロセスを配置してブロック
    します。S--;
}
atomic Signal (Semaphore S)
{ S++;
}
```

ミューテックスは、0か1の値しか持たない2値セマフォに過ぎません。セマフォは一般的なロックであり、制限はありません（つまり、ミューテックスが2つの値しか持たないのに対し、一般的なセマフォはそうではありません）。また、上記のコードにはatomicキーワードがあることにも注目してください。これは、コードが実行されても決して中断されないことを意味しています。つまり、1つのプロセッサ上でコードのブロックとして正しい順序で実行されることが保証されているのです。これらは1つのユニットとして扱われることになっています（そのため、アトミックオペレーションと呼ばれています）。

残念ながら、上で示したような単純なものではありません。アトミック操作はハードウェアに依存しているため、動作させるためにはプロセッサの支援が必要です。具体的には、LOCK命令のプレフィックスを利用する必要があります。この点については、後にこれらのプリミティブを実際のコードに実装する際に、より詳しく説明します。

今のところ、セマフォの使用例を見るのが一番だと思います。セマフォは最初に導入するときには難しいかもしれません。マルチプロセッシングを完全にサポートするためには、セマフォを頻繁に使用することになるので、セマフォの使い方を練習することが重要です。

Process A	count++; signal (s);
Process B	wait (s); count--; signal (s);

スピンドル

相互排除は、クリティカルセクション問題を解決するための最初の基準です。これは、あるプロセスがクリティカルセクションに入ると、他のプロセスはクリティカルセクションに入ることができないということです。この機能を実装するためには、相互排除を保証できるようなアトミックな操作を実装する何らかの方法が必要になります。1つのアイデアは、単純な変数をロックの役割として使用することです。ロックが1であれば、あるプロセスがクリティカルセクションの内部にいることになります。そこで、最初のアイデアは

```
int lock=0;
```

Process A	Process B
while(1) { if (!lock) lock = 1; do_something(); lock=0; }	while(1) { if (!lock) lock = 1; do_something(); lock=0; }

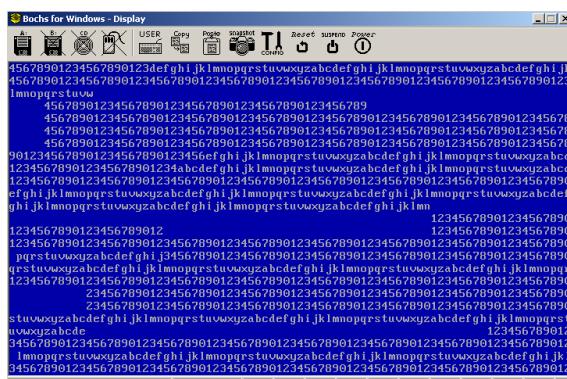
かなりシンプルですね。ロックは0から始まるので、最初に実行されたプロセスはこれを検出してロックを設定します。それが終わるとロックを解除し、2番目のプロセスがそれを使えるようにします。これははある程度うまくいきますが、まだ大きな問題があります。例えば、プロセスAがロックが0であることを検出したが、プロセスAがロックを設定する前にプロセスBによって中断されたとします。そこで、プロセスBはロックが0であることも検出し、今度はロックを設定します。そのため、プロセスBがdo_somethingのどこかで中断されても、プロセスAはロックが0のままであるかのように実行を続けます。そのため、ロック変数を読み込んでロックしようとしてプロセスが中断されても、両方のプロセスが同時に同じクリティカルセクション（この例では、クリティカルセクションはdo_somethingの呼び出し）に入ることができます。これは小さなエラーのように見えますが、すぐに伝播してしまい、かなりの頻度で発生してしまいます。

つまり、ここでの問題は、ロックへのアクセスと設定が、中断されることなく行えることを保証できないということです。この操作はアトミックではありません。

実際にアトミックな操作をしなくても何が起こるかをイメージできるように、2つのスレッドがあるとします。1つ目のスレッドは文字のa-zを表示し、2つ目のスレッドは数字の0-9を表示します。これらのスレッドは、後に開発するスケジューラを使って同時に実行されます。以下がそのスレッドです。

Process A	Process B
<pre>void task_1() { char c='a'; while(1) { DebugPutc(c++); if (c>'z') c='a'; } }</pre>	<pre>void task_1() { char c='0'; while(1) { DebugPutc(c++); if (c>'9') c='0'; } }</pre>

この2つのタスクが並行して実行されると、出力はインターリープされた混乱状態になります。理由は、両方のプロセスが共有リソースを無順序に読み書きしているからです。先に述べたようにロックを導入したとしても、出力はあまり良くなりません。この例では、共有リソースはビデオメモリと、カーソルの位置決めとスクロールを担当するDebugPutcが使用するグローバル変数です。あるプロセスが現在のxやyの位置を読み取ったり、スクロールの準備をしていると、そのプロセスが中断され、最初のプロセスが知らないうちに位置やその他のグローバル変数が変更されてしまう可能性があります。



セマフォを使わないサンプル。出力がめちゃくちゃになっていることに注意してください。

この問題を解決するためには、単純なロックだけではなく、何かが必要です。私たちの方向性は良いのですが、ハードウェアのサポートが必要です。もし、ロック変数のテストと設定を1回の操作で行うことができる方法があれば、絶対に中断されないことが保証されるので（つまりアトミック）、最終的には相互排除基準を満たすことができます。

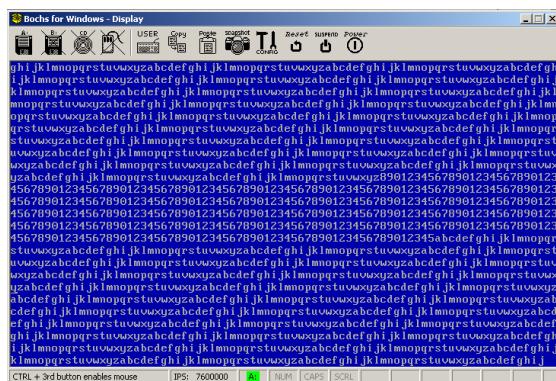
そのようなハードウェアプリミティブの一つが、LOCK命令プレフィックスです。このプレフィックスは、その命令が実行されている間、システムバスをリード/ライトからロックします。データバスがロックされているので、アトミック性が保証されています。そのため、ロック変数の設定やテストを行う際には、単純にLOCK XCHG や LOCK BTS を使用することができます。例えば、以下のようになります。

```
inline void acquire(int* lock) {
    _asm{
        mov eax,[lock];
        A:    [EAX], 0 ポーズ
        jc a
    }
}
inline void release(int* lock) {
    _asm{
        mov eax,[lock]
        mov [eax], 0
    }
}
```

これで、これらの関数を呼び出して、ロックの獲得と解放ができるようになりました。

Process A	Process B
<pre>void task_1() { char c='a'; while(1) { acquire(lock); DebugPutc(c++); release(lock); if (c>'z') c='a'; } }</pre>	<pre>void task_2() { char c='a'; while(1) { acquire(lock); DebugPutc(c++); release(lock); if (c>'z') c='a'; } }</pre>

そして望み通りの結果を得ることができる。



スピノロックを使ったサンプルの実行。表示がきれいに整ったことに注目してください。

4. Classic Concurrency Problems

生産者・消費者問題（境界線付きバッファ問題）

これは、最初に見ていく古典的な同時実行問題です。2つの独立したプロセスがあるとします。1つはプロデューサー、もう1つはコンシューマと呼ばれます。また、両方のプロセスが使用している共有バッファがあるとします。プロデューサーはバッファにデータを入れる役割を果たし、コンシューマーはデータを取り出す役割を果たします。これが、古典的な「プロデューサー/コンシューマー問題」(Bounded Buffer Problem) の基本的な設定です。問題は、プロデューサーが、バッファがすでにいっぱいになっている場合にデータを追加しないようにすることと、コンシューマーが、空のバッファからデータを取り出そうとしないようにすることです。この問題は、複数のプロデューサーとコンシューマーがいる場合に、より興味深いものとなります。

Example. This is a solution to the Bounded Buffer problem. This assumes a single producer and consumer running concurrently.

```
Semaphore c = 0;  
Semaphore s = BUFFER_SIZE;
```

```

Producer           Consumer
while (true) {      while (true)
    { item = produce ();   wait(c);
        wait(s);
        write(item);
        signal(c);
    }
}

```

読者と作家の問題

典型的なリーダー/ライター問題は、あるオブジェクトが多くのプロセスで共有され、リーダーとライターという2種類のプロセスが存在する場合です。読者は共有データを読みますが、変更することはありません。ライターはデータを読み、変更することができます。多くのリーダーが同時にデータを読むことができます。

Example. There are many different solutions and versions of this problem, this is one of them. Note that we use two semaphores here so that we can allow multiple readers at the same time.

```

Semaphore c = 1;
Semaphore s = 1;
int count = 0;

Writer           Reader
while (true) {   while(true) {
    wait(c);      wait(s);
    write();       count++;
    signal(c);    if (count == 0)
}                   wait(c);
                     signal(s);
                     read();
                     wait(s);
                     count--;
                     if (count == 0)
                         signal(c);
                     signal(s);
}

```

5. Inter-Process Communication

IPC (Inter-Process Communication) とは、OSがサポートする技術で、プロセスが他の実行中のプロセスと信号やデータを共有することを可能にする。IPCプロトコルの実装にはさまざまな種類の技術があるが、ここではよく使われるものを紹介する。

パイプ

パイプは、プロデューサーとコンシューマーの間でデータを保存するために円形のバッファを使用する基本的な技術です。プロデューサはバッファにデータを書き込み、コンシューマはバッファからデータを読み出します。データのプロデューサーとコンシューマーは複数存在することができます。パイプには、匿名パイプと名前付きパイプの2種類があります。名前付きパイプには名前が付けられ、仮想ファイルシステムのファイルオブジェクトとして表示されます。システム内のどのプロセスでも名前付きパイプを開くことができます。匿名パイプは、親プロセスから継承した子プロセスのみが開くことができます。

オペレーティングシステムは、コンシューマとプロデューサの間で共有されるデータストリームを格納する機能、ストリームを読み書きする機能、読み取るべきデータがないときにパイプから読み取ろうとするプロセスをブロックする機能を提供する必要があります。オペレーティングシステムは、上で説明した相互排除技術を使って、読み書きを同期させなければなりません。これは通常、先入れ先出し (FIFO) の円形バッファを使用し、セマフォを使用して読み書き時のアクセスを同期させます。

パイプはファイルシステムのオブジェクトです。パイプを開くと、ファイル記述子のポインタが戻ってきます。そのため、ファイルのReadおよびWriteメソッドを使って、パイプをファイルのように読み書きすることができます。開いたファイルハンドルは子プロセスに継承されますので、パイプも継承されます。

パイプは、ファイルシステムの記述子と同様に管理することができます。プロセス・パラメータ・ブロックには、ファイル・ディスクリプターやパイプなどのシステム・オブジェクトへのオープンな参照をすべて格納したプロセス・ハンドル・テーブルへのポインタが格納されています。また、デバイスファイルをすでにサポートしているシステムでは、些細なことでも実装することができます。

メッセージの受け渡し

基本的には、プロデューサーがメッセージを送信し、コンシューマーがそれを受信するというシンプルなアイデアです。しかし、メッセージの受け渡しを同期で行うか、非同期で行うかによって、さらに問題が出てきます。また、メッセージをどのように保存するか、どこで管理するか、メッセージのフォーマット、メッセージが期待されたフォーマットで期待されたプロセスに配信されているかどうかをどのように検証するかなども考えなければなりません。

では、メッセージとは一体何なのでしょうか？メッセージは、プロセスが望むものなら何でもあります。コンシューマーとプロデューサーは、メッセージをどのように解釈するかについて、何らかのプロトコルに同意する必要があります。両者はメッセージのデータ構造を知っている必要があります。オペレーティングシステム側では、マイクロカーネルのようにOSが定義したメッセージでない限り、OSはデータのフォーマットを気にしません。

オペレーティングシステムには、最低限、メッセージの送受信をサポートする機能が実装されている必要があります。

同期型メッセージパッシング

同期メッセージパッシングには、最低でも2つの関数が必要です。JとKをプロセス識別子(PID)とすると

```
send(J, message)
receive(K, &message)
```

producerはsendを呼び出してメッセージを投稿します。同期メッセージパッシングでは、producerは、Jがreceiveを呼び出してメッセージを取得するまで、中断されたキューリングに入れられます。Jがreceiveを呼び出すと、オペレーティングシステムは、Jに送られたメッセージを直接コピーし、Jを再開することができます。その後、オペレーティングシステムは、プロデューサーを待ち行列に戻し、スケジューラが実行できるようにします。同期メッセージ・パッシングでは、同時に実行されるのは一方（プロデューサまたはコンシューマ）だけなので、メッセージ・キューは必要ありません（他方はサスペンションまたはウェイティングになります）。

Asynchronous Message Passing

非同期のメッセージパッシングにも最低2つの関数が必要です。

```
send(J, message)
receive(K, &message)
```

プロデューサは send を呼び出してメッセージをポストし、コンシューマは receive を呼び出してメッセージを取得します。非同期メッセージパッシングでは、OSがプロセスごとにメッセージキューを管理します。プロデューサはいつでもメッセージを送信でき、中断されることはありません。メッセージはメッセージキューの最後にコピーされます。その後、消費者はメッセージキューの先頭からメッセージを受け取ることができます。メッセージキュー自体はカーネルメモリ内に割り当てられ、プロセス制御ブロック内の専用ポインタがキューを指し示します。

ここで、興味深い問題があります。同期メッセージパッシングでは、プロセスがreceiveを呼び出したときに、メッセージを送信したプロセスがないと、他のプロセスがsendを呼び出すまでプロセスが中断してしまいます。非同期のメッセージパッシングでは、2つの選択肢があります。

受信やを呼び出したプロセスを中断することができます。
受信側にステータスコードを返させて、現在のプロセスを継続させることができます。

やはり、もう少し機能を充実させた方が良いのではないかでしょうか。

```
send(process, message)
receive(process, &message)
sendrec(process, &message)
notify(process, message)
```

共有メモリ

同じ物理フレームを2つ以上のプロセスの仮想アドレス空間にマッピングすると、そのプロセス間で共有されます。両方のプロセスが同じページを読み書きできるようになります（ページをマッピングする際に設定されたセキュリティ属性によります。（例えば、物理フレームを、プロセスAでは読み取り/書き込み、プロセスBでは読み取り専用としてマッピングすることができます）。オペレーティングシステムでは、一般的にメモリマップドファイルによって共有メモリをサポートしています。例えば、Windowsでは、まず、名前の付いたメモリマップドファイルオブジェクトのCreateFileまたはOpenFileを呼び出し、続いて、メモリの領域をプロセスアドレス空間にマップし、そのポインタを返すMapViewOfFileを呼び出します。

6. Scheduling

1. スケジューラーは、システムリソースの割り当てを行います。システムリソースには、CPU、メモリ、システムデバイスなどがあります。スケジューラーは一般的に多数存在しますが、短期、中期、長期の3つのカテゴリーに分類される傾向があります。
2. **Long term schedulers** are responsible for admitting processes into the system and terminating them.
3. **Medium term schedulers** are responsible for suspending and resuming processes.
4. **Short term schedulers** are responsible for allocating CPU time and dispatching processes.

このセクションでは、マルチタスク・システムを実装するためのコア・コンポーネントである短期スケジューラについて主に説明します。そこで、ここでのデモの目標は、短期スケジューラを作成することです。

スケジューリングアルゴリズム

使用できるアルゴリズムには様々なものがあり、中には複雑なものもあります。ここでは、一般的なアルゴリズムを紹介しますが、デモをシンプルにするため、ラウンドロビン方式を採用します。

First Come First Serve

FCFS (First Come First Serve) では、ジョブは来た順に実行されます。アルゴリズムはその名の通りシンプルで、スケジューラーが最初のジョブを選択して実行させます。次に2番目のジョブ。次に2番目のジョブ、そして次のジョブ.....という具合です。このアルゴリズムは、Ready Queueに入っているジョブを、入ってきた順に循環させます。前のジョブが終了するまで、新しいジョブは開始されません。このアルゴリズムは、ブリエンプティブなマルチタスクにはあまり適していません。

Example. In the following example, P1 arrives at time 0, P2 arrives at time 1, and P3 arrives at time 2. These processes are placed in the **Ready queue** to be executed. P1 is the first job, so the algorithm selects it to be run. P2 is selected next, but only after P1 is completed. P2 does not get selected until time=5.

Process	Arrive	Run time	Service time
P1	0	5	0
P2	1	3	5
P3	2	8	8

Shortest Job First

SJF (Shortest Job First) アルゴリズムでは、各ジョブの実行に必要な時間をシステムが知る方法が必要です。このアルゴリズムでは、Ready Queueから次に実行されるジョブのうち、時間の差分が最も小さいものを選択します。このアルゴリズムには、プロセスの飢餓の問題があります。時間差が小さいジョブが優先されると、ジョブがReady Queueに残ってしまうことがあります。この例は、前述のFCFSアルゴリズムと非常によく似ていますが、時間差を計算する必要があるため、実際にはほとんど実装されていません（プロセスが実行される時間を事前に知るために、ソフトウェアがオラクルである必要があります）ので、別の例は必要ないと思います。

Priority Queue

このシステムでは、各ジョブに優先度の高い番号を割り当てることができます。優先度の高いジョブが先に選択される。これが優先スケジューリングアルゴリズムの基本的な考え方です。優先度をどのように決定するかは設計者の自由です。同様に、2つの優先度が同じ場合にどのように処理するかも設計者次第である。ひとつのアイデアは、デフォルトの優先順位を設け、それをユーザーが調整できるようにすることです。2つの優先度が同じ場合は、FCFSやSJFを使ってどちらを使うかを決めます。もう一つのアイデアは、プロトコルに基づいて優先順位を計算することです。プロトコルは、システム管理者が割り当てる、システムのリソースやメモリの制約を測定して算出したりします。後述するように、優先順位は他のスケジューリングアルゴリズムと一緒に使われることが多いです。

要約すると、Ready Queueから優先度の高いジョブを選択すればよいのです。このアルゴリズムは、SJFと同様、優先度の高いプロセスが優先度の低いプロセスを駆逐してしまうため、プロセスの飢餓状態に陥ります。

Round Robin

システムは、各プロセスに量子と呼ばれる実行するためのタイムスライスを与えます。その後、システムは現在実行中のプロセスを先に実行し、別のプロセスの実行を許可します。プロセスは、レディキューに表示された順に選択されます。すべてのプロセスの実行が許可されるため、このアルゴリズムはプロセスを飢えさせることはありません。システムは、実行するように選択されたプロセスの実行状態を保存・復元するために、コンテキストスワッピングを行います。コンテキストスワッピングについては、後ほどマルチタスクについて説明する際に取り上げます。

Example. Given processes P1, P2, P3 and a time quantum of 5, the **Round Robin (RR)** algorithm first selects P1 to run. After the quantum time is up, the system **preempts** P1. P1 is moved to the back of the **Ready Queue**. The system saves the context of P1. The system then selects P2 and the system performs a **context switch**. P2 can now execute.

Process	P1	P2	P3	P1	P2	P3
Quantum=5	0	5		10	15	20

Multilevel Queue

次に何を実行するかを決めるのに、1つのReady Queueを使う代わりに、複数のReady Queueを使ってはどうでしょうか。特権レベルと別のスケジューリングアルゴリズムを、マルチレベルキューに組み合わせることで、両方の世界を手に入れることができるというものです。

基本的な考え方は、複数のキューを持つことです。そして、これらのキューは異なる優先度のためのものです。例えば、5つの優先度がある場合、5つのキューを持つことになります。アルゴリズムはまず、優先度の高いキューから、優先度に基づいて実行するジョブを選択します。そのキューに複数のジョブがある場合は、別のアルゴリズム（RRなど）を使って実行するジョブを決定します。また、異なる優先度のキューに対して、異なるスケジューリングアルゴリズムを使用することもできます。このアルゴリズムは、優先度の高いスケジューリングと同じ理由で、プロセスを飢えさせてしまう可能性があります。このように、素晴らしいアルゴリズムがありますが、プロセスのスタービングを防ぐにはどうしたらよいでしょうか？

Multilevel Feedback Queue

Multilevel Feedback Queueは、プロセスの飢餓を防ぐために、マルチレベルキューを改良したものです。マルチレベルキューの問題点は、ある優先度LのプロセスがキューLに挿入されたとき、優先度がLよりも大きい新しいジョブを投入するだけで、プロセスを飢えさせてしまうことでした。つまり、ある優先度のキューから別の優先度のキューにプロセスを移動させることができます。

上記の例では、優先度Lのプロセスは、ある程度の時間が経過すると、より高い優先度のキューに移動します。これは、プロセスが最高の優先度のキューに到達するまで続きます。したがって、プロセスが飢餓状態に陥ることはありません。また、重要なシステムタスクを実行する必要がある場合に役立つかかもしれない、より低い優先度のキューに移動させることで、ジョブの優先度を下げることもできます。マルチレベルのフィードバックキューを実装する際の難しさは、いつプロセスを移動させるべきかを決定することです。これは、今日の最新のオペレーティングシステムで使用されている最も一般的なアルゴリズムです。

Example. The following is an example of a **multilevel queue**. Here we have three queues, system processes have the highest priority and applications have the lowest priority. Different scheduling algorithms can be used on each of the different queues to select jobs from them. The scheduler selects the highest priority **non-empty** queue. It then uses another algorithm (such as **FCFS** or **RR**) to select a job from that queue. In **multilevel feedback queues**, the system can move processes between different queues. For example, we can move jobs from L3 then L2 then L1 over time, thereby raising its priority so it can run. Thus no process starvation.

Queue Level	Priority	Queue
L1	System	Processes
L2		Batch Jobs
L3		Applications

7. Multitasking

この章では、これまで多くの内容を取り上げてきました。そしてようやく、この章のメインとなる部分にたどり着くことができます。
マルチタスクである。すべてをコードにまとめていきます。

最初にプロセスの状態管理を取り上げたのは、スケジューラやマルチタスクコンポーネントがプロセスを選択し、異なる状態間で移動させる必要があるからです。例えば、スケジューラはプロセスをReadyからRunningに切り替える必要があります。より高度なページング技術（ページ・スワッピング・アルゴリズムなど）をサポートする予定であれば、Suspended状態との間でプロセスを切り替えられるようにする必要があります。システムは、Suspendedプロセスと、メモリ内で完了信号を待っているプロセスとを区別できる必要があります。どちらのプロセスもPCB (Process Control Block)を持ち、システムリソースを使用していますが、Suspendedプロセスはメモリを使用していません。また、プロセスを一時停止する方法が必要でした。そのため、Waitステートを導入しました。ご覧のように、状態管理はマルチタスクを実装するための重要なコンポーネントです。そのため、最初にこれを取り上げました。

次に見たのは「プロセス作成」です。状態管理でどのように使われるのかを詳しく調べました。第24章では、CreateProcess関数を実装しました。この関数では、PE (Portable Executable) イメージをメモリにロードし、仮想アドレス空間にマッピングし、ユーザー modeで実行したことを見出してください。このセクションでは、この関数を使って新しいプロセスを作成し、スケジューラが選択するReadyキューに追加していきます。

続いて、コンカレント・プログラミングの入門編を見てみましょう。クリティカルセクション問題」「相互排除」「セマフォ」などのトピックを取り上げました。並行処理とは、複数のプロセスやスレッドが非同期に実行されることです。並行プログラミングは、非同期のプロセス間の通信を同期させるための技術を提供します。並行プログラミングは難しく、正しい方法はありません。並行処理を使用すると、コードには必ずバグがあります。この章のテーマがマルチタスクであることから、コンカレント・プログラミングを紹介しました。共有リソースはマルチタスクと密接な関係があるので（一般的には共有ライブラリ、シグナル、メッセージパッシングなどの形で）、ここでは簡単な紹介をしました。

続いて、IPC (Inter-Process Communication) の紹介を行いました。IPCは、シンプルなOS以外では重要な役割を果たします。また、マルチタスクでIPCをサポートするシステムには、本章で説明するコンカレント・プログラミング技術が必要です。システムコールを使ってIPCの一形態をすでに使用しています。

最後にスケジューリングアルゴリズムを取り上げました。スケジューラーは、オペレーティングシステムの心臓部です。実行するプロセスを選択する役割を担っており、マルチタスクシステムの中核となるアルゴリズムです。

1. さて、いよいよ今回は、マルチタスクOSの世界に飛び込んで、まとめていきます。ご存知のように、マルチタスクには3つのタイプがあります。
 2. Preemptive
 3. Non-Preemptive
 4. Cooperative

ここでは、「プリエンプティブ・マルチタスク」に注目します。

プラン

1. 今回は、ラウンドロビン（RR）というスケジューリングアルゴリズムを使用します。このアルゴリズムでは、選択されるプロセスにリソースとして量子を割り当てられることができます。そこで必要になるのがクロックです。システムには様々な種類のクロックがあります。

2. Programmable Interval Timer (PIT)
3. Advanced Programmable Interrupt Controller (APIC) timer
4. Real Time Clock (RTC)
5. High Performance Event Timer (HPET)
6. etc.

今回のデモでは、PITがすでにサポートされているため、PITを使用することにします。これで、使用するスケジューリングアルゴリズムとクロックが決まりました。第24章では、PCB (Process Control Block) とTCB (Thread Control Block) を紹介しました。今回はTCBを拡張して、現在のスレッドの状態を保存したり、ユーザー モードからカーネル モードへの切り替えに必要な情報を追加します。

```
typedef struct _thread
{
    uint32_t esp;
    uint32_t ss;
    uint32_t kernelEsp;
    uint32_t kernelSs;
    struct _process* parent;
    uint32_t priority;
    int state;
    ktime_t sleepTimeDelta;
} thread;
```

スレッドに関連するタスクを作成するには、いくつかの低レベルのものが需要です。スタックは、現在のレジスタコンテキストを格納します。上の構造体のespフィールドに指定されたスタックに、レジスタコンテキストを格納します。スケジューラは、タスクの作成、タスクの管理、およびタスクの切り替えを行います。以下のセクションでは、それぞれの機能を詳しく説明します。なお、サンプルコードはすべて本章最後のデモプログラムで使用しています。

レディ キュー

まず、これらのタスクを格納する場所が必要です。タスクは、カーネルのメモリアロケータによって、ページングされていないプールから動的に割り当てられるべきです。しかし、シリーズにはカーネルアロケータが実装されていないので、実装には配列を使うしかありません。ラウンドロビン・スケジューリングに必要な先入れ先出しの機能は、循環型の待ち行列を使って実装することができます。これは、キューの一番上の要素を削除して後ろに追いやるだけで、次のタスクに移行できるというものです。つまり、新しいタスクがキューの一番上になります。

```
スレッド _readyQueue[THREAD_MAX]; int _queueLast, _queueFirst;スレッド _idleThread;スレッド *_currentTask;スレッド _currentThreadLocal;

/* キューをクリアします。*/
void clear_queue() {
    _queueFirst = 0;
    _queueLast = 0;
}

/* スレッドを挿入します。*/
bool queue_insert(thread t) {
    _readyQueue[_queueLast % THREAD_MAX] = t;
    _queueLast++;
    return true;
}

/* スレッドを削除します。*/
thread queue_remove() {
    thread t = _readyQueue[_queueFirst % THREAD_MAX];
    _queueFirst++;
    return t;
}

/* キューの先頭を取得します。*/
thread queue_get() {
    return _readyQueue[_queueFirst % THREAD_MAX];
}
```

今回の例では、準備の整ったタスクのための単一のキューを実装するだけです。タスクは、キューをシャッフルすることで、いつでも削除、追加することができます。ここで、_currentTask ポインタに注目してください。第 25 章では、このポインタは常に _currentThreadLocal を指し、現在実行中のスレッドのローカルコピーを保存します。ISR はこのポインタを使ってスレッドの状態を保存したり復元したりします。次のセクションでは、このISR

について説明します。

割り込みサービスルーチン (ISR) について

さて、最初の課題は、タイマーがトリガーされたときにスケジューラーを呼び出すことです。ハードウェア割り込みは、割り込みコントローラ（ここではレガシーのProgrammable Interrupt Controller (PIC)）によって発生します。もちろん、マルチプロセッサ(MP)やCPU間のIRQに使用されるAPIC(Advanced PIC)などもありますが、本シリーズではシンプルにするためにレガシーPICインターフェースのみをサポートしています。PICは、PITから送られてくるIR#0信号のように、ハードウェアデバイスがPICに信号を送ると、CPUに信号を上げます。そして、PICは別の信号（この場合はCPUのIRQライン）を上げてCPUに通知します。どのIRQが呼び出されるかは、PICをどのようにプログラムしたかによります。IR#0をISR33にマッピングするようにPICをプログラムしたことを思い出してください。これが意味するところは、PITが発火するたびに、CPUは現在のコードの実行を停止し、リターンcs、eip、フラグを現在のスタックにプッシュしてから、IDT (Interrupt Descriptor Table) にインストールしたISR、つまりIDT[33]を呼び出すということです。

つまり、タイマーISRはすでに割り込みベクター33にインストールされています。これはプロテクトモードの設定時に行いました。これは、ハードウェア割り込みを有効にするために必要でした。それはそれでいいのですが、私たちがしたいのはそれを上書きすることです。

これを実現するのが割り込みチャイニングです。前の章で割り込みチャイニングを紹介しましたが、実際に実践することはありませんでした。今までは、です。必要なのは、古いISRを取得して、独自のISRをインストールすることです。今すぐ実行しましょう。

```
/* register isr */ old_isr =
getvect(32);

setvect (32, scheduler_isr, 0x80);

簡単ですね。IDTの話をしたときにgetvectとsetvectを実装しました。それをIDT[32]にインストールしたのは、そこにPIT ISRがあったからです。これは、old_isrに保存して、新しいISRであるscheduler_isrをインストールするためです。
```

以上のことを考慮して、PITが起動するたびに、代わりにscheduler_isrが呼ばれることになります。さて、次は難しい部分、つまりISRの記述です。ISRが何をする必要があるのか、いつ呼び出されるのかを考えてみましょう。ISRはいつでも呼び出すことができます。しかし、タスクが実行されているときは常に呼び出されます。必要なのは、現在のレジスタの状態を保存し、スケジューラを呼び出すことだけです。PICにEOI (End-Of-Interrupt) を送ることも忘れずに。

まず、今回のデモのために実装されたISRを紹介し、以下ではそれが何をしているのかを詳細に説明するために、パートごとに分解していきます。

```
_declspec(naked) void _cdecl scheduler_isr () {
    __asm {
        ; 割り込みを解除してコンテキストを保存します。
        ; クリップ
        ; シュド
        ; 現在のタスクがない場合は、単にリターンします。
        mov eax, [_currentTask]
        cmp eax, 0
        jz interrupt_return
        ; セレクターを保存します。
        ; プッシ
        ; ウ ds プ
        ; ツ シュ
        es プッ
        ; シュ fs
        ; プッシ
        ; ウ gs
        ; カーネルセグメントに切り替えます。
        ; mov ax, 0x10
        ; mov ds, ax
        ; mov es, ax
        ; mov fs, ax
        ; mov gs, ax
        ; SAVE STATION
        ; mov eax, [_currentTask]
        ; mov [eax], esp
        ; 呼び出しスケジューラー。
        ; call scheduler_tick
        ; を復元します。
        ; mov eax, [_currentTask]
        ; mov esp, [eax].
        ; tss_set_stack(kernelSS, kernelESP)を呼び出します。
        ; このコードは、後でユーザーのタスクで必要になります。
        ; push dword ptr [eax+8]
        ; push dword ptr [eax+12]
        ; call tss_set_stack
        ; esp, 8を追加
        ; EOIを送信して文脈を復元します。
        ; ポップ
        ; G、ポ
        ; ップF、
        ; ポップ
        ; ES、ポ
        ; ップ
        ; DS
        interrupt_return
    }
}
```

```
す。  
: 古いISRを呼び出す必要があるかどうかをテストします。  
: mov eax, old_isr  
: cmp eax, 0  
: jne chain_interrupt  
:  
old_isrがnullの場合、EOIを送信して戻る。
```

```

;
mov al,0x20
out 0x20,al
popad
iretd
;
old_isrが有効であれば、それにジャンプする。これは
PITタイマーの割り込みハンドラを作成します。
;
chain_interrupts
    ポバド
    jmp old_isr
}

```

ISRは、現在のレジスタ・コンテキストの保存と、現在のタスクのスタック・ポインタの保存を行います。その後、スケジューラを呼び出して、現在のタスクのスタック・ポインタを復元し、前に保存したレジスタ・コンテキストを復元します。すべてが復元されているので、ISRが戻ってきたときには、タスクは問題なく実行され続けています。ISRは一見複雑に見えますが、実際にはそうではありません。もう少し詳しく見てみましょう。他のISRと同様に、まず最初に行なうことは、現在のレジスタの状態を保存してスタックに保持することです。つまり、ISRは次のように始まります。

```
_declspec(naked) void _cdecl scheduler_isr () {
```

```
    _asm { cli
        pushad

```

```
        ポバドア
        イレド
    }
```

PITによってインストールされたISRの上にISRをインストールするので、ここでは細心の注意を払う必要があります。これは、scheduler_isrがクロックティックごとに呼び出されることを意味します。setvectを呼び出してインストールすると、レディキューにタスクがないうちにPITが起動してしまうことがあります。実行するタスクがないときは、何もすることがないので、ISRを返してほしいだけです。また、割り込みを無効にしても、元に戻さないことに気づくかもしれません。これは問題ありません。現在実行中のタスクは、FLAGSレジスタを通じて割り込みを有効にします。FLAGSレジスタはすべてのケースで保存されているので、IRETDを発行すると、リターン時にFLAGS.IFが有効になり、割り込みが再び有効になります。私たちのISRは次のようになります。

```
_declspec(naked) void _cdecl scheduler_isr () {
```

```
    _asm { cli
        pushad

```

```
        ;
        現在のタスクがない場合は、単にリターンします。
        ;

```

```
        mov eax, [_currentTask]
        cmp eax, 0

```

```
        jz interrupt_return

```

```
        ;
        ;<実際のISRコードはこちら
        ;

```

```
interrupt_returnです。
    ポバドア
    イレド
}
```

最後に、PITハードウェアがscheduler_isrを呼び出しているため、PITドライバのISRは決して呼び出されていないことを覚えておく必要があります。私たちは割り込みを連鎖させたいのです。つまり、先にインストールされていた古いISRがあれば、それを実行するチャンスを与えたいのです。これは、そのISRにジャンプする（呼び出さない）ことで行われます。別のISRを呼び出すときには、そのISRが別の割り込みを連鎖させるか、連鎖を断ち切るためにEOI（End-Of-Interrupt）コマンドを発行することを念頭に置く必要があります。別のISRを呼び出す場合、技術的にはまだ割り込みを処理しているので、EOIを送信する必要はなく、IRETDも必要ありません。しかし、別のISRを呼び出さず、元のプロセスに制御を戻す場合は、両方が必要です。つまり、私たちのISRは次のようになります。

```
_declspec(naked) void _cdecl scheduler_isr () {
```

```
    _asm {

```

```
        ;
        割り込みを解除してコンテキストを保存します。
        ;

```

```
        クリップ
        シュド

```

```
        ;
        現在のタスクがない場合は、単にリターンします。
        ;

```

```
        mov eax, [_currentTask]
        cmp eax, 0

```

```
        jz interrupt_return

```

```
        ;
        ;<実際のISRコードはこちら
        ;

```

```
interrupt_returnで
す。

```

```
        ;
        古いISRを呼び出す必要があるかどうかをテストします。
        ;

```

```
        mov eax, old_isr
        cmp eax, 0

```

```
        jne chain_interrupt

```

```
        ;
        old_isrがnullの場合、EOIを送信して戻る。
        ;

```

```
        mov al,0x20

```

```
        out 0x20,al

```

```
        popad

```

```
        iretd

```

```
        ;

```

```
        old_isrが有効であれば、それにジャンプする。これは

```

```

PITタイマーの割り込みハンドラを作成します。
;
chain_interrupt:
    ポパド
    jmp old_isr
}

```

実際にタスクを実行するISRの本体は、以下の部分です。

```

; セレクターを保存します。
;
push ds
push es
push fs
push gs
;
カーネルセグメントに切り替えます。
;
mov ax, 0x10
mov ds, ax
mov es, ax
mov fs, ax
mov gs, ax
;
SAVE STATION
;
mov eax, [_currentTask]
mov [eax], esp
;
呼び出しスケジューラー。
;
call scheduler_tick
;
を復元します。
;
mov eax, [_currentTask]
mov esp, [eax].
;
;tss_set_stack (kernelSS, kernelESP)を呼び出します。
;このコードは、後でユーザーのタスクで必要になります。
;
push dword ptr [eax+8]
push dword ptr [eax+12]
call tss_set_stack
esp, 8を追加
;
;restore context.
;
pop bp
G、 pop
pop f,
pop pp
ES、 pop
pop p
DS

```

まず、セグメントレジスターをスタックにプッシュします。(この前にPUSHADを行ったことを思い出して下さい。また、CPUはISRが最初に呼ばれたときにCS、EIP、EFLAGSもスタックにプッシュしました)。これらをスタックに格納することで、現在のスレッドレジスタのコンテキストを保存することができます。これらのレジスタがスタック上にプッシュされた順番は、後にstackFrame構造体で使用する順番と一致します。次に、これらのセグメントレジスターを、グローバルディスクリプタテーブル (GDT) からずっと前に設定したカーネルモードセレクタに設定します。このようにするのは、現在実行中のタスクがカーネルモードのタスクであるという仮定をしていないからです。もし、タスクがユーザーモードのタスクであれば、DS、ES、FS、GSは0x10ではなく0x23のままで。スレッドスタックに元のタスクセレクタを保存しておいたので、今からでも調整できます。ユーザーモードタスクから来た場合、CPUはタスクステートセグメント (TSS) からSSとCSを自動的に設定しますので、これらはすでに適切に設定されています。スタックについては、後でもう少し詳しく見てみましょう。最後に、ESPの現在の値を_currentTask->espに保存し、scheduler_tickを呼び出します。

_currentTaskは、ISRによって、現在実行中のタスクを常に指していると見なされます。スケジューラがタスクを変更した場合、その新しいタスクが「現在」実行中の新しいタスクになります。新しいタスクであっても、ESPをその新しいタスクに戻すだけです。

_currentTask->esp フィールドを使用しています。最初にレジスタコンテキストを新しいスレッドスタックに保存したので、それらをそれぞれのレジスターに戻すだけです。ずっと前に実装したtss_set_stackも呼びます。これは、復帰するタスクがユーザーモードタスクの場合にのみ有効です。新しいタスクのカーネルスタックを更新することで、TSSにセットします。今度のデモでは、それぞれが1つのカーネルスタックを持つカーネルスレッドのみを実行するので、これはまだ適用されません。しかし、ユーザーレベルのスレッドは、ユーザー空間とカーネル空間の両方で実行されるため、1つではなく2つのスタックを持っていることを覚えておいてください。次のいくつかの章では、アドレス空間の管理とユーザ空間について詳しく説明します。

では、どうやってタスクを切り替えるのでしょうか？スケジューラが呼び出されたときに _currentTask ポインタが変化したらどうなるかを少し考えてみましょう。この新しいタスクのレジスタ・コンテキストとスタック・ポインタは同じように保存されているので、scheduler_tick関数の内部でこのポインタを変更するだけで、ISRは自動的に新しいタスクのレジスタ・コンテキストとスタックをロードします。このように、タスクの切り替えは、そのポインタを更新するだけで簡単にできます。

タスクの切り替え

つまり、タスクの切り替えには、ポインタの更新が必要なのです。ラウンドロビン・スケジューリングでは、実行中のタスクを格納するためにキューを使うことができます。キューはすでに先入れ先出しで動作しているので、現在のタスクを削除して再挿入するだけで、タスクをプッシュバックすることができます。これにより、コードが大幅に簡素化されます。

/* 次のタスクをスケジューリ

```
ングします。 */ void dispatch
() {
    /* ここではラウンドロビンを行い、削除と挿入を行います。
注意: _currentTask ポインタは、常に
_currentThreadLocal. そのため、_currentThreadLocal を更新するだけです。
*/
    /* queue_remove();
queue_insert(_currentThreadLocal);
_currentThreadLocal = queue_get(); */
}
/* 時計の目盛りごとに呼び出されます。 */
```

```
void scheduler_tick () {
```

```
    /* ディスパッチャーを実行するだけです。*/ dispatch();  
}
```

以上が、その内容です。上記はラウンドロビン・スケジューリングを実装しており、一定の量子が経過するとタスクを入れ替えます。タスクは先ほど実装したReady Queueに格納されます。これで残るのは、タスクの作成です。

上記は複数のスレッドに対して動作しますが、異なるプロセスに属するスレッドに対しては動作しません。典型的な解決策は、現在のスレッドの親プロセスと新しいスレッドを比較することです。両者が同じプロセスに属していれば、ディスパッチャは単にリターンすることができます。両者が異なるプロセスに属している場合、ディスパッチャはVMMを呼び出して新しいプロセスのアドレス空間に切り替える必要があります。サンプルコードをシンプルにするために、この章ではこれを避けることにしました。しかし、次の章でアドレス空間の管理について詳しく説明する際には、これをサポートする予定です。

タスク作成

例えば、スケジュール関数が _currentTask ポインタを別のタスクに更新したとします。そこで、この関数がISRに戻ると、ISRはIRETDを発行する前に、この新しいタスクからスタックとレジスタのコンテキストを設定します。これはうまくいきますが、タスクがすでにスタックとレジスタ・コンテキストをスタック上に持っている場合に限ります。

そのため、最初にタスクを作成する際に設定する必要があります。そこで、基本的なスタックフレームを設定し、タスクのespとeipをスタックとエントリーポイントの関数に設定します。スタックフレームは、私たちのISRが期待しているものでなければなりません。ISRに戻ると、まずPOP GS、POP FS、POP ES、POP DSを行い、次にPUSHAを行い、続いてIRETDを行います。PUSHAはEAX, EBX, ECX, EDX, ESI, EDI, ESP, EBPをポップします。そしてIRETDはEIP, CS, FLAGSをポップします。つまり、これがタスクが生成されたときの初期スタックフレームになるはずです。

```
typedef struct _stackFrame { uint32_t  
gs;  
uint32_t fs;  
uint32_t es;  
uint32_t ds;  
uint32_t eax;  
uint32_t ebx;  
uint32_t ecx;  
uint32_t edx;  
uint32_t esi;  
uint32_t edi;  
uint32_t esp;  
uint32_t ebp;  
uint32_t eip;  
uint32_t cs;  
uint32_t flags.  
} stackFrame;
```

```
タスク task_create (uint32_t entry, uint32_t esp) { スレッ  
ド t;  
stackFrame* frame = ((stackFrame*) esp);  
frame->flags = 0x202;  
frame->cs = 8;  
frame->eip = (uint32_t)entry;  
frame->ebp = 0;  
frame->esp = 0;  
frame->edi = 0;  
frame->esi = 0;  
frame->edx = 0;  
frame->ecx = 0;  
frame->ebx = 0;  
frame->eax = 0;  
frame->ds = 0x10;  
frame->es = 0x10;  
frame->fs = 0x10;  
frame->gs = 0x10;  
t.esp = (uint32_t) frame;  
t.ss = 0x10;  
を返します。  
}
```

これはほとんどのタスクで動作しますが、1つだけ例外があります - 初期タスクです。今回作成したISRは、現在実行中のコードがタスク内にある場合にのみ動作します。これも鶏と卵の問題です。これを回避するためには、特別なタスクオブジェクトを作成し、マルチタスクを開始する準備ができたらそれを実行する必要があります。

```
static thread _idleTask;
```

```
void task_execute(thread t) { ...
```

```
_asm{  
    mov esp, t.esp  
    pop gs  
    ポップfs  
    ポップes  
    ポップds  
    ポップad  
    iretd  
}
```

```
/* スケジューラを初期化します。*/  
void scheduler_initialize(void) { ...
```

```
    /* 準備完了したキューをクリアします。*/ clear_queue();
```

```
/* プロセス・リストをクリア
します。*/
/* init_process_list(); */

/* アイドル・スレッドを作成して追加します。*/
_idleThread = thread_create(idle_task, (uint32_t)create_kernel_stack(), true);
/* 現在のスレッドをアイドル・タスクに設定して追加する。*/
_currentThreadLocal = _idleThread;
_currentTask          = &_currentThreadLocal;
```

```

queue_insert(_idleThread);

/* register isr */ old_isr =
getvect(32);
setvect (32, scheduler_isr, 0x80);
}

/* アイドルタスク */
void idle_task() {
    while(1) _asm pause;
}

```

上記はすべてをまとめています。アイドルタスクを作成し、キューに追加し、ISRをインストールして、初期タスクを実行します。初期タスクが実行されると、PITが起動するたびにISRが呼び出され、必要に応じてスケジューラーを呼び出して現在のタスクを更新します。

8. Introduction to MP

ここでは、他のプロセッサを起動するための標準的なインターフェイスとIPI (Inter-Processor Interrupt) を提供するために設計されたMP (Multi-Processor) 仕様について、ごく簡単に紹介します。これは、コンカレント・プログラミングの難易度を一気に高めることができるため、上級者向けのトピックだと考えています。当社のスケジューラは一度に1つのタスクしか実行しませんが、MPでは独立したCPUのスケジューリングを行う低レベルのスケジューラを実装することができ、複数のタスクを同時に実行することができます。MP規格の詳細を知りたい方は、MP仕様書をご覧になることをお勧めします。お使いのシステムが、MPで使用されるIOAPIC、LAPIC、ICIをすでにサポートしている必要があります。

マルチプロセッシングには対称型マルチプロセッシング (SMP) と非対称型マルチプロセッシング (ASP) がある。SMPではすべてのプロセッサが同じ種類であるのに対し、ASPでは同じ種類ではない。デスクトップシステムではASP方式は非常に珍しいため、ほとんどのシステムはSMPのみをサポートしている。しかし、MP規格はどちらにも対応しており、拡張性の余地があるため、より多様なマシンタイプに採用することができ、さらにOSがさまざまなタイプのシステムに対応して設定できるようになっている。

システムが最初に起動するとき、ハードウェアはBoot-Strap Processor (BSP) を選択し、起動する唯一のプロセッサとして機能します。BSPは、最初に起動するプロセッサであり、最後にシャットダウンするプロセッサでなければなりません。オペレーティングシステムは、BSPから他のアプリケーションプロセッサ (AP) にSTARTUP IPIを送信し、APを起動させることができます。他のAPは、BSPでも他のAPでも起動することができます。STARTUP IPI (およびINIT IPI) は、オペレーティングシステムが他のプロセッサを起こすために送るものです。

オペレーティングシステムは、システムがMPをサポートしているかどうかを検出するために、まず、フローティングMPフローティングポインタ構造体を検索する必要があります。この構造体には、MPコンフィギュレーションテーブルの物理アドレスが含まれている。コンフィギュレーション・テーブルは読み取り専用である。このテーブルには、ローカルAPIC (LAPC) のメモリマップドアドレス、プロセッサエントリ (プロセッサLAPIC IDを含む)、IOAPICエントリ (IOAPICベースのメモリマップドアドレスを含む)、バス、および割り込みコンフィギュレーションエントリが格納されています。オペレーティングシステムは、BSPのLAPIC IDを記憶して、BSPが最後にシャットダウンされるようにしなければなりません。

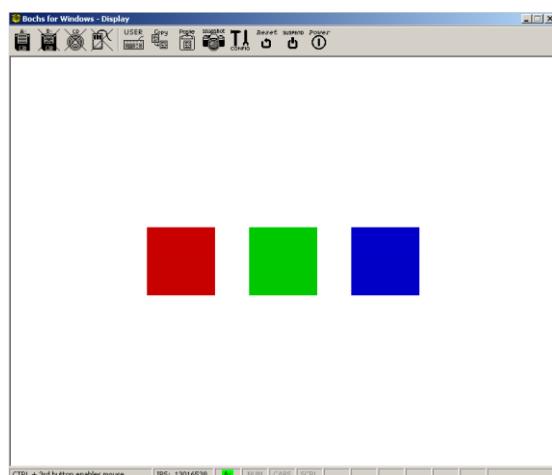
別のAPをウェイクするためには、BSP LAPICまたは別のAP LAPICを介してINIT IPIを送信すればよいです。LAPICのメモリマップされたレジスタは、MPコンフィギュレーションテーブルのプロセッサ情報に格納されています。その後、そのAPにSTARTUP IPIを送り、実行を開始する必要があります。本当にそれだけのことなのです。INIT IPIは、APをリセットさせます。STARTUP IPIは、リアルモードで指示された場所で実行を開始させます。オペレーティングシステムは、BSPで行ったように、APIをプロテクトモードやロングモードで構成するためのリアルモードスタブルーチンを提供する必要があります。

以上、マルチプロセッサ・システムの簡単な紹介をしました。他のプロセッサ（またはプロセッサコア）を起動するのは非常に簡単で、スケジューラを実装した後にSMPを試してみることをお勧めします。MPについては、APICを使った後のチュートリアルで詳しく説明する予定です。MPに興味のある方のために、その概要と方向性をお伝えしたいと思います。

9. Demo

[デモをダウンロードする]をクリックしてください。

新しいコードのほとんどは上記の文章で説明されており、私たちは最初のリリースを準備しているだけです。ストレステストと最終的な統合で問題が発生しなければ、デモは来週か再来週中にリリースされる予定です。



3つのタスクを実行する800x600x32モードのデモ。

このデモは、私たちにとって初めての本格的なグラフィカル・デモです。このデモでは、3つのタスクを同時に実行し、それぞれのタスクが実行中にビデオメモリの色を選択して循環することで、実行中であることを視覚的に示しています。テキストベースのデモではなく、グラフィカルなデモを選んだのは、次のような理由からです。

グラフィックスシリーズを読んでいなくても大丈夫です。グラフィックスシリーズに目を通していなくても、ここで説明します。

Bochs Graphics Adapter (BGA)について

この章の主要なテーマに集中できるように、コードをできるだけシンプルにするため、システムがISA用に構成されていることを前提にBGAを使用することにしました。このコードはBochs固有のもので、実際のシステムでは動作しません。実際のシステムでは、PCIバスのインフラをスキャンする必要がありますが、これはより高度な章のトピックになるかもしれません。

```
#define VBE_DISPI_IOPORT_INDEX          0x01CE
#define VBE_DISPI_IOPORT_DATA           0x01CF
#define VBE_DISPI_INDEX_XRES            0x1
#define VBE_DISPI_INDEX_YRES            0x2
#define VBE_DISPI_INDEX_BPP             0x3
#define VBE_DISPI_INDEX_ENABLE          0x4
#define VBE_DISPI_DISABLED              0x00
#define VBE_DISPI_ENABLED               0x01
#define VBE_DISPI_LFB_ENABLED           0x40

void VbeBochsWrite(uint16_t index, uint16_t value)
{ outportw (VBE_DISPI_IOPORT_INDEX, index);
  outportw (VBE_DISPI_IOPORT_DATA, value);
}

void VbeBochsSetMode (uint16_t xres, uint16_t yres, uint16_t bpp)
{ VbeBochsWrite (VBE_DISPI_INDEX_ENABLE, VBE_DISPI_DISABLED);
  VbeBochsWrite (VBE_DISPI_INDEX_XRES, xres);
  VbeBochsWrite (VBE_DISPI_INDEX_YRES, yres);
  VbeBochsWrite (VBE_DISPI_INDEX_BPP, bpp) と
    なります。
  VbeBochsWrite (VBE_DISPI_INDEX_ENABLE, VBE_DISPI_ENABLED | VBE_DISPI_LFB_ENABLED);
}
```

ビデオモードの設定は、VbeBochsSetModeを呼び出すだけです。この例では、800x600x32がよくサポートされているようなので、800x600x32を使用しています。ISAのリニアフレームバッファ(LFB)は、あらかじめ定義された0xe0000000の位置にあります。しかし、ここではページングを有効にしているので、LFBを使用するためには、仮想アドレス空間にマッピングする必要があります。今回のデモでは、0x200000仮想アドレスにマッピングします。このマッピングは、LFBのサイズをページ数で計算し、VMMを呼び出して各ページをマッピングします。

void* VbeBochsMapLFB () { ... }

```
/* BGAのLFBはISAシステムではLFB_PHYSICALになります。*/
#define LFB_PHYSICAL 0xE0000000
#define LFB_VIRTUAL 0x200000

/* LFBを現在のプロセスのアドレス空間にマッピングする。
 * int pfcnt =
 * WIDTH*HEIGHT*BYTES_PER_PIXEL/4096; int c;
 for (c = 0;c <= pfcnt; c++)
  vmmngr_mapPhysicalAddress (vmmngr_get_directory(),LFB_VIRTUAL + c * 0x1000,LFB_PHYSICAL + c * 0x1000, 3) を実行します。
 /* LFBへのポインタを返します。*/
 return (void*) LFB_VIRTUAL;
}
```

上記の関数により、0x200000に書き込むことで、LFBに描画することができます。ディスプレイ上のゴミを掃除するために、次にそれをクリアします。大量のピクセルを描画する必要があるので、32ビットモード用に関数を最適化するようにしています。この関数は画面を白にします。

```
void fillScreen32 () {
  uint32_t* lfb = (uint32_t*) LFB_VIRTUAL; for
  (uint32_t c=0; c<WIDTH*HEIGHT; c++)
    lfb[c] = 0xffffffff;
}
```

32 Bits Per Pixelモードでは、ピクセルカラーは、赤8ビット、緑8ビット、青8ビットで構成されます。上位8ビットは今回の目的では無視されますが、通常は透明度の値として使用されます。ここでは3つのタスクを使って、3つの長方形をレンダリングし、3つの色の強度を循環させています。ここでは、ディスプレイ上の異なる場所にレンダリングするので、同時実行性の問題を心配する必要はありません。ディスプレイのメモリは共有されていますが、各タスクは別々の部分にレンダリングします。

```
void rect32 (int x, int y, int w, int h, int col) { uint32_t* lfb =
  (uint32_t*) LFB_VIRTUAL;
  for (uint32_t k = 0; k < h; k++) for
  (uint32_t j = 0; j < w; j++)
    lfb[(j+x) + (k+y) * WIDTH] = col;
}
```

void kthread_1() { int col = 0; bool dir = true; while(1) { rect32(200, 250, 100, 100, col << 16); if (dir){ if (col++ == 0xfe) dir=false; }else if (col-- == 1) dir=true; } }	void kthread_2 () { int col = 0; bool dir = true; while(1) { rect32(350, 250, 100, 100, col << 8); if (dir){ if (col++ == 0xfe) dir=false; }else if (col-- == 1) dir=true; } }	void kthread_3 () { int col = 0; bool dir = true; while(1) { rect32(500, 250, 100, 100, col); if (dir){ if (col++ == 0xfe) dir=false; }else if (col-- == 1) dir=true; } }
--	--	---

スレッドスタック

通常、スレッドは2つの独立したスタックを持っています。ユーザー modeで実行するときのスタックと、カーネル modeで実行するときのスタックです。スレッドがユーザー modeで実行されているとき、CPUはタスクステートセグメント (TSS) のesp0およびss0フィールドを取得することで、カーネルスタックに切り替えることを思い出してください。スケジューラは、TSSを新しいスレッドのカーネル modeスタックに更新する役割を担っています。ただ

し、第25章については、すべてのスレッドがカーネル空間で実行されるため、TSSが参照されることはありません。言い換えれば、第25章のスレッドは、カーネルモードスタックという1つのスタックしか持ちません。

次の2つの章でアドレス空間の管理について説明しますが、その中でユーザー mode のスレッドをサポートする予定です。将来のアドレス空間アロケータを使用して、各ユーザモードスレッドのためにユーザ空間にスタック空間を確保します。これは、スレッドがユーザモードとカーネルモードの両方のスタックを持つことを意味します。

スレッドは、CPL (Current Privilege Level) が0のコードを実行する際に、カーネルモードスタックを使用します。CPLがTSSからのRPL (Requested Privilege Level) よりも小さい場合、CPUは自動的にこれをロードします。つまり、ユーザモードのスレッドが実行されていて、PITが発火したとします。すると、CPUはSS=TSS.ss0、ESP=TSS.esp0を設定します。そして、リターンCSとIPをこの新しいスタックにプッシュし、ISRを呼び出します。ISRが終了すると、IRETを実行してユーザモードのコードとスタックに戻ります。

このため、ユーザーレベルのスレッドは、最低でも2つの独立したスタックを持たなければなりません。最初のスタックはカーネル空間にマップされていなければならず、もう一つのスタックはユーザ空間にマップされていなければならず、プログラムが実行中にアクセスできなければなりません。カーネルレベルのスレッドは1つのスタックしか必要ありません。

アドレス空間アロケータがないので、まだユーザモードのスタックをうまく割り当てることができず、(ハックしないと) ユーザーレベルのスレッドをサポートできません。また、まだ適切なカーネルモードアロケータがないので、カーネルレベルのスタックの割り当てもうまくサポートできません。これらは次の章か2章のトピックになります。

そこで、第25章では、カーネルメモリに領域を確保し、4kブロックごとにスタックを割り当てるにしました。

```
void* create_kernel_stack()
{
    physical_addr      p;
    virtual_addr       location;
    void*              ret;

    /* 4kのカーネルスタック用にこの領域を確保している。 */ #define
    KERNEL_STACK_ALLOC_BASE 0xe0000000

    /* スタック用に4kフレームを確保する。 */ p =
    (physical_addr) pmmngr_alloc_block(); if (!p)
    return 0;
    /* 次の空き4kメモリブロック */

    location = KERNEL_STACK_ALLOC_BASE + _kernel_stack_index * PAGE_SIZE;
    /* カーネル空間にマッピングします。 */
    vmmngr_mapPhysicalAddress (vmmngr_get_directory(), location, p, 3);

    /* スタックの一番上を返します。 */ ret =
    (void*) (location + PAGE_SIZE);
    /* 再び呼ばれたら次の4kを確保する準備をする。 */
    _kernel_stack_index++;

    /* そして、スタックの先頭を返す。 */
    return ret;
}
```

Back to Sleep()

フロッピーディスクの読み取り動作を遅らせるために、非常に基本的なスリープ機能を実装したこと思い出してください。この実装では、単にビギー・ループに入りて時間を浪費していました。今度はそれをスレッドシステムに採用してみましょう。

基本的な考え方は、sleepは関数を呼び出したスレッドを一時停止させることです。つまり、現在のスレッドの状態をREADYからBLOCKに調整し、タスクスイッチを強制的に行う必要があります。スケジューラはブロックされたスレッドを追跡し、適切に処理する必要があります。これは通常、他のオペレーティングシステム・コンポーネントからのシグナルを介して行われます。例えば、あるスレッドがデバイスの準備が整うのを待っている場合、そのスレッドはブロックされることがあります。ここでシステムは、そのスレッドがドライバからシグナルを受け取るまで待つ必要があります。それまでは、スケジューラは他のスレッドの実行に移るべきです。デモを比較的シンプルにするために、私たちは少し違ったやり方を選びました。

必要なのは、現在実行中のプログラムの状態を変更し、タスクスイッチを強制的に行うことです(int 33を介してISRを直接呼び出すことで)。スケジューラには、ブロックされたスレッドをチェックしながら、実行する新しいスレッドを選択するためのロジックコードが含まれます。次のスレッドがブロックされている場合は、そのスリープ時間のデルタをデクリメントし、スリープ時間のデルタがゼロになった時点でスレッドを目覚めさせます。

このデモではsleepを使用していませんが、ディスクドライバのコードはsleepに依存しています。これで、ディスク・デバイスからの読み取りを試みるスレッドは、適切にスリープできるようになりました。

メインプログラム

最後に、メインプログラムを見てみましょう。第25章のデモでは、スタックをカーネル空間に移動し、ブートローダから渡されたブートパラメータブロックの静的なコピーを作成した後、スタックを再調整しました。その後、上述のサービスを使用して、ビデオモードの設定、スケジューラの初期化、3つのスレッドの作成とレディキューへの追加を行いました。スレッドはカーネル空間で動作するため、カーネルスタックのみが割り当てられており、create_kernel_stackを呼び出して割り当てます。

また、第25章で作成したスレッドシステムと互換性を持たせるために、第24章のプロセス生成・管理コードを完全に書き換えました。ただし、次の章で行うユーザモードスタックの割り当てをサポートするまでは完成しません。

```
void __cdecl kmain (multiboot_info* bootinfo) {...
```

```
/* カーネルサイズを保存し、ブートインフォをコピーします。*/
_asm mov    word ptr [kernelSize], dx
memcpy(&_bootinfo, bootinfo, sizeof(multiboot_info));
/* スタックの調整 */
_asm lea esp, dword ptr [_kernel_stack+8096] init
(&_bootinfo);

/* ビデオモードの設定とフレームバッファのマッピング
を行う。 */ VbeBochsSetMode(WIDTH,HEIGHT,BPP);
VbeBochsMapLFB();
fillScreen32 () です。
```

```

/* スケジューラを初期化します。*/
/* scheduler_initialize(); */
/* カーネル・スレッドを作成します。*/
queue_insert (thread_create(kthread_1, (uint32_t) create_kernel_stack(),true));
queue_insert (thread_create(kthread_2, (uint32_t) create_kernel_stack(),true));
queue_insert (thread_create(kthread_3, (uint32_t) create_kernel_stack(),true));

/* アイドル・スレッドを実行します。*/
/* execute_idle(); */

/* これは決して実行されてはいけません。
* for (;;) __asm {cli
    hlt}となります。
}

```

10. Conclusion

この章では、スケジューリング・アルゴリズム、SMPの概要、コンカレント・プログラミング、そしてプリエンプティブなラウンドロビン・スケジューラの実装について説明しました。また、BDA(Bochs Graphics Adapter)を使った高解像度ビデオモードの紹介や、状態管理、IPC技術の紹介も行いました。

次の章では、いよいよカーネルやユーザモードのアロケータ、アドレス空間の割り当て、ページスワッピング、ページフォルトの処理などのメモリ割り当てアルゴリズムについて説明します。この章では、フリーリストとスタックアロケータ、SLABアロケータ（およびその変種）、ZoneとArenaアロケータ、Buddyアロケータ、ユーザ空間の管理、再帰的ページディレクトリ、ページファイルとスワップ空間、およびその他のトピックを取り上げます。この章の内容を発展させて、ユーザモードのプロセスローディングをサポートしていきます。膨大な量の資料が出てくるため、この章は1つまたは2つの別々の章になるかもしれません。

次の機会まで。

~Mike () です。

OS開発シリーズ編集部

[Home](#)



オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - Portable Executable (PE)

by Mike, 2011

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

Introduction

Welcome!

よーし、これは長くなりそうだ。

この章では、上級者向けのトピックであるPE実行ファイル形式を取り上げます。PEリソース、ダイナミックリンクなどのカバーを見ていきます。また、この章は、情報を可能な限り完全なものにするために、より多くの情報を含むアップデートを予定しています。

本章で紹介する内容のほとんどは情報提供のみを目的としており、完全性を保つためと、読者がサポートを提供したいと思う場合にのみ掲載しています。また、これらの情報の多くは、公式のPE仕様書にも記載されていますのでご注意ください。

この章が終われば、ローダーを開発してシングルタスク環境をサポートするために必要なものがすべて揃うことになります。

さあ、始めましょう。

File Format

Abstract

PE (Portable Executable) ファイル形式は、WindowsやReactOSなどのWindowsライクなOSで使用されている標準的な実行ファイル形式です。また、EFI (Extensible Firmware Interface) マシンでの起動に使用される標準的なファイル形式でもあります。

PEの実行ファイルフォーマットは、再配置、シンボルテーブル、リソース、ダイナミックバイニングなどをサポートする複雑なフォーマットです。

Terms

VA (バーチャルアドレス)

仮想アドレス (VA) は、現在のプログラムの仮想アドレス空間 (VAS) 内の線形アドレスです。PEの実行形式では、すべてのアドレスが仮想アドレスです。これらのアドレスは32ビットのリニアアドレスです。

RVA (相対的仮想アドレス)

RVA (Relative Virtual Address) とは、実行プログラムのベースアドレスからの相対的なVAのことです。PEの実行形式では多くの部分でRVAが使用されていますので、RVAとは何か、そしてRVAからリニアアドレスを得る方法を知っておくことが重要です。RVAはベースアドレスからの単なるオフセットでしかありません。ですから、リニアアドレスを得るには、ベースにRVAを加えるだけです。

```
Linear address = Base address + RVA
```

解析時に多くの箇所でこの計算を行う必要があるため、この点は重要です。

セクションとセクションテーブル

セクション

高度な実行ファイルフォーマットでは、リンクプロセスを簡素化し、ソフトウェアに構造を与るために、プログラムセクションを使用するのが一般的です。セクションは、実行イメージやオブジェクトファイルに格納される命令やデータの標準的な方法を提供することで、リンクプロセスを簡素化します。

セクションには通常、その中にどのような要素があるかに関連した名前があります。例えば、「.data」は、変数や初期化されていないデータを含む一般的なセクション名です。他のセクション名には、歴史的な背景があります。例えば、.textは、実行可能なコードやオブジェクトコードを含むセクションの典型的な名前です。.bssは、一般的にグローバルな、プログラム全体で初期化されたデータに使用されます。

C++ツールチェインを例にとると、グローバルな名前空間や静的に定義された変数は.bssに格納されます。また、コンパイル後に生成されるバイトコードは.textに格納されます。

PEの実行ファイル形式には、通常、以下のセクション名が1つ以上含まれています。

- .テキスト
- .データ
- .bss
- .アーチ
- .edata
- .idata
- .pdata
- .rdata
- .reloc
- .rsrc

.sbss
.sdata
.srodata
.xdata

セクションテーブル

プログラムファイルやオブジェクトファイルには、複数のセクションが含まれています。各セクションの基本位置とセクション名は、通常、セクションテーブルに格納されている。セクションテーブルは、単純な構造体のリンクリストであったり、ハッシュテーブルであったり、さまざまな形式があります。

シンボルとシンボルテーブル

シンボルマーク

C++でプログラミングをしていると、有名な「未定義シンボル」のリンクエラー（Cの実装によっては警告）に遭遇することが多いと思います。その通り、完全にエラーなしでコンパイルとリンクを行うことができます。）これは、関数を呼び出したり、変数を参照する際に、リンクの段階で定義が解決できなかったために起ります。

関数や変数は、リンカーによってシンボルと呼ばれます。シンボルには、名前のほか、データ型や値などの情報が含まれています。コンパイル時には、最終的なプログラムをリンクできるようにするために、これらのシンボルを追跡する必要があります。もし、現在の翻訳ユニットで定義されていないシンボルが使用されていて、それがEXTERNシンボルである場合、コンパイル担当者はシンボルをオブジェクト・ファイルに書き込む際に、EXTERNシンボルとしてマークする必要があります。

リンクの段階で、EXTERNとマークされたシンボルがまだ値を持っていない（シンボルが定義されていない）場合、リンカーは上記のエラーを発行します。

シンボルは、モジュール、翻訳ユニット、またはライブラリ間で変数や関数を定義するためのものです。

プログラム全体、およびプログラムがリンクしているライブラリ全体で、同じ名前のシンボルは1つしかありません。このため、高級言語を使用すると名前が衝突する可能性が高いため、変数名や関数名には通常、名前のマングリングが適用されます。もちろん、これはアセンブリ言語には適用されません。適用される名前のマングリングは複数の要因に依存し、ツールチェーンによって異なります。

ちょっと見てみましょう。ここでは、いくつかのC関数の宣言と、その右にマングルされたシンボル名を示しています。マングルされた名前の中の数字は、パラメタのバイト数です。

```
void __cdecl function (int i);    -> _function
void __stdcall function(int i);   -> _function@4
void __fastcall function(int i);  -> @function@4
```

また、__cdeclの呼び出し規則を持つ関数には、前にアンダースコアが付いているだけです。これにより、アセンブリ言語を使ってC言語の関数を簡単に定義することができ、Cコードでその関数を簡単に呼び出すことができます。

C++の名前の付け方には標準がありません。あるコンパイラーでは ?h@@YAXH@Z のようなシンボリックな名前を生成しますが、他のコンパイラーでは

7h Fi@W?h\$n(i)vは、void h(int)という同じ機能のためのものです。このため、アセンブリ言語での使用は現実的ではありません。しかし、それでも可能である。

記号表

セクションテーブルと同じように、シンボルテーブルがあります。シンボルテーブルは、ソフトウェアがシンボル名とシンボルに関する情報（エクスポートされたシンボルであるかどうか、データタイプ、プロパティなど）を検索するための手段です。シンボルテーブルは通常、情報のリンクリスト、またはハッシュテーブルで実装されています。

Structure

アブストラクト

MSVC++の章で、PE実行形式の構造を見てきました。PE実行ファイルをメモリにロードすると、そのメモリには、ロードしたファイルの正確なコピーが格納されます。つまり、PEファイルフォーマットの最初の構造内の最初のバイトは、実際にファイルがメモリにロードされた場所からの最初のバイトに位置しています。

例えば、PEファイルを1MBにロードした場合、インメモリのフットプリントは以下のようになります。



上の画像は、MSVCの章を読んだことのある読者には見覚えがあるはずです。上の図を見ると、PEファイルが1MBまで読み込まれていた場合、ディスク上の最初の構造体であるIMAGE_DOS_HEADERがメモリ上のその位置から始まり、その後にファイル内の残りの構造体（パディングを含む）が続きます。

上の画像も単純化しすぎっていて、決してPEファイルフォーマットの全体像を示しているわけではありません。PEファイルフォーマットの構造はかなり大きく、多くの構造体やテーブルで構成されています。

1. ここでは、完全なフォーマットをご紹介します。

2. IMAGE_DOS_HEADER structure (**Important**)
3. STUB program
4. IMAGE_FILE_HEADER structure [COFF Header] (**Important**)
5. IMAGE_OPTIONAL_HEADER structure (**Important**)
6. Segment Table
7. Resource Table
8. Resident Name Table
9. Module Reference Table
10. Imported Names Table
11. Entry Table
12. Non Resident Name Table
13. Segments
 1. Data
 2. Info

上の表は、ファイル形式の最初から最後までを示しています。重要と書かれた項目は、プログラムを実行するために解析方法を知る必要があります。その他の情報は、情報提供のみを目的としています。IMAGE_OPTIONAL_HEADER 構造体で重要なメンバは、エントリポイントのアドレスを含むメンバと、イメージベースアドレスだけです。

このファイルの各セクションの解析については、次のセクションで詳しく説明します。また、テーブルやディレクトリを解析する際に使われる他の構造についても同様に紹介します。

IMAGE_DOS_HEADER 構造体

IMAGE_DOS_HEADERは、PEファイルの最初の構造体です。この構造体には、プログラムファイルとそのロード方法に関するグローバルな情報が含まれています。この構造体に含まれる情報のほとんどは、DOSソフトウェアに関連するものであり、後方互換性のためにのみサポートされています。

構造はフォーマットに沿っています。

さてさて、この構造には面白いものがたくさんあります。初期CS:IPと初期SS:SPのメンバーは、OSが通常CSのためにスタックスペースとコード記述子の値を割り当てているので、無視してください。これらのメンバーは、DOS領域やv8086モードを必要とするソフトウェアの時に顕著でした。

STUBプログラム

さて、それでは。IMAGE_DOS_HEADER構造体の直後にDOSスタブプログラムがあることに注目してください。これ、実は便利なプログラムなんです。これは、DOSの中からWindowsのプログラムを実行しようすると、「このプログラムはDOSモードでは実行できません」と表示するプログラムです。

リンクオプションの/STUBを使ってスタブプログラムを変更することができます。

```
/stub=myprog.exe
```

DOSが実行ファイルをロードしようとすると、DOSはIMAGE_DOS_HEADER構造を解析し、有効なDOSプログラムであるため、DOSスタブプログラムを実行しようとします。Win32サブシステムで実行する場合、Windowsローダはスタブプログラムを無視します。

image_nt_headers

STUBプログラムの後には、PEヘッダー構造体のフォーマットを含む、IMAGE_NT_HEADERSという構造体があります。以下がその構造体です。

記号は「PE¥0¥0」と一致しなければなりません（¥0はヌル文字）。IMAGE_FILE_HEADERには、ローダが使用する追加情報と、IMAGE_OPTIONAL_HEADER構造体のサイズがすべて含まれています。IMAGE_OPTIONAL_HEADERは、ファイルの中で一番大きくて重要な構造体です。また、サイズは定義されています。

この構造体の位置を特定するために、OSローダーはIMAGE_DOS_HEADERのe_lfanewメンバを使用する必要があります。e_lfanewはメモリ上のこの構造体への RVAなので、この構造体の位置を特定するために、ローダーは次のように実行する必要があります。

```
IMAGE_DOS_HEADER* pFile = (IMAGE_DOS_HEADER*) imageBase;
IMAGE_NT_HEADERS* pHeaders = (IMAGE_NT_HEADERS*) (pFile->e_lfanew + imageBase);
```

これはimageBaseがプログラムファイルがメモリに読み込まれた場所を参照していると仮定しています。DOSなどの古いOSでは、このヘッダーのメンバーを認識していないので、これらのOSでは無視されます。

この構造体は、他の2つのヘッダー構造体のフォーマットを含んでいます。では、最初の構造体を見てみましょう。

イメージファイルのヘッダー

IMAGE_FILE_HEADERは、COFF (Common Object File Format) のヘッダー構造です。以下のような形式になっています。

この構造はあまり複雑ではありません。上記のほとんどは、デバッガにとってのみ有用です（シンボルテーブルの解析）。SizeOfOptionalHeaderは重要です。

IMAGE_OPTIONAL_HEADERはサイズが定義されていないので、このメンバで構造体のサイズを知ることができます。

Machineは以下の値のいずれかです。x86マシン

の場合は0x014c

0x0200 (x64マシン用) 0x8664

(AMD64マシン用)

通常のケースでは、x86アーキテクチャ用に開発しているため、0x014cとなるはずです。

特徴は、リンカーがビットごとにORすることができるビットフラグで構成されており、ローダーに実行可能なイメージの種類に関するさまざまな特性を知らせることができます。以下はそのフォーマットです。

- ビット0：セットされている場合、画像は再配置情報を持たない。
- ビット1：セットされている場合、ファイルは実行可能です。
- ビット2：セットされている場合、画像にはCOFFライン番号がありません。
- ビット3：セットされている場合、画像にはCOFFシンボルテーブルのエントリがありません。
- ビット4：セットされている場合、イメージの作業セットをトリムする。(Windowsのメモリ管理に依存します。廃止されました)
- ビット5：セットされている場合、ローダーは実行ファイルが2GB以上のVAを扱えると仮定します。
- ビット6：セットされている場合、ローダーはイメージが32ビットワードをサポートしていると仮定します。
- ビット7：セットされている場合、画像はデバッグ情報を持たない
- Bit 8: 設定されている場合、イメージはネットワークドライブから直接実行されません(Windows特有)。
- ビット9：設定されている場合、画像はSYSTEMファイルとして扱われます。
- ビット10：セットされている場合、イメージはDLLファイルとして扱われる
- ビット11：セットされている場合、イメージはシングルプロセッサのマシンでのみ実行されます。
- ビット12：セットされている場合、ピッグエンディアン。obsoleteフラグ

Windowsのヘッダには、IMAGE_FILE_RELOCS_STRIPPEDやIMAGE_FILE_EXECUTABLE_IMAGEなどの定数が定義されており、これらのフラグを設定する際に使用することができます。

ご覧のように、この構造のほとんどは、画像をどのようにロードするかというローダーの情報のみのためのものです。でも、待ってください。リソース、シンボルテーブル、デバッグ情報などはどうなっているのでしょうか？IMAGE_OPTIONAL_HEADERにサイズが定義されていない理由が見えてきました。ちょっと見てみましょう。

image_optional_header

うわ、出てきた。これは、このファイルの中で最も複雑な構造です。良いニュースは、この構造を見たことがあるということです。

```

USHORT Magic;                                // not-so-magical number
UCHAR MajorLinkerVersion;                    // linker version
UCHAR MinorLinkerVersion;
ULONG  SizeOfCode;                           // size of .text in bytes
ULONG  SizeOfInitializedData;                // size of .bss (and others) in bytes
ULONG  SizeOfUninitializedData;              // size of .data,.sdata etc in bytes
ULONG  AddressOfEntryPoint;                  // RVA of entry point
ULONG  BaseOfCode;                          // base of .text
ULONG  BaseOfData;                          // base of .data
ULONG  ImageBase;                           // image base VA
ULONG  SectionAlignment;                   // file section alignment
ULONG  FileAlignment;                      // file alignment
USHORT MajorOperatingSystemVersion;         // Windows specific. OS version required to run image
USHORT MinorOperatingSystemVersion;
USHORT MajorImageVersion;                  // version of program
USHORT MinorImageVersion;
USHORT MajorSubsystemVersion;               // Windows specific. Version of SubSystem
USHORT MinorSubsystemVersion;
ULONG  Reserved1;
ULONG  SizeOfImage;                         // size of image in bytes
ULONG  SizeOfHeaders;                       // size of headers (and stub program) in bytes
ULONG  CheckSum;                            // checksum
USHORT Subsystem;                          // Windows specific. subsystem type
USHORT DllCharacteristics;
ULONG  SizeOfStackReserve;                  // size of stack, in bytes
ULONG  SizeOfStackCommit;                  // size of stack to commit
ULONG  SizeOfHeapReserve;                  // size of heap, in bytes
ULONG  SizeOfHeapCommit;                   // size of heap to commit
ULONG  LoaderFlags;                        // no longer used
ULONG  NumberOfRvaAndSizes;                // number of DataDirectory entries

```

まず、最後のメンバーであるDataDirectoryを見てみましょう。定数である IMAGE_NUMBEROF_DIRECTORY_ENTRIES は、長年にわたって変更されてきました。これは、この構造体のサイズを変えることができるメンバーです。このメンバーについては、後ほど詳しく見てみましょう。

明らかにオプショナルではないのに、なぜ「オプショナル」ヘッダーと呼ばれているのか気になりますよね。これは、COFFオブジェクトファイルではオプションとなっているからです。実行イメージではオプショナルではありませんが、オブジェクトファイルではオプショナルです：)

マジックは以下のいずれかになります。

- 0x10b : 32 ビットの実行形式
- 0x20b : 64 ビットの実行形式
- 0x107 ROMイメージ

通常の場合は、0x10bになるはずです。

この構造のメンバーの多くは、それほど複雑ではありません。

サブシステム・メンバーは、Windows固有のものです。プログラムを正しく実行するために必要なサブシステムをWindowsに伝えます。以下の値のいずれかになります

す（ここでは完全性を期して掲載しています）。

- 0: 不明
- 1: ネイティブサブシステム
- 2 : GUIサブシステム
- 3: CUIサブシステム
- 5 : OS/2 CUIサブシステム
- 7 : POSIX CUIサブシステム
- 9 : Windows CE GUIサブシステム
- 10 : EFI
- 11 : EFIブート ドライバ
- 12 : EFIランタイム ドラ
イバ
- 13 : EFI ROM

16: ブートアプリケーション

DllCharacteristicsメンバは、ローダにDLLに関する情報を与えるビットフラグを含んでいます。以下のような形式になっています。

ビット0-3: 予約済み

ビット4: セットされている場合、DLLは再配置可能である

ビット5: セットされている場合、強制的にコードインテグリティチェックを行う

ビット6: セットされている場合、画像はDEP (Data Execution Prevention) に対応しています。

ビット7: セットされている場合、画像を分離してはならない

ビット8: セットされている場合、画像は構造化された例外処理 (SEH) を使用しま

せん ビット9: セットされている場合、画像はバインドされません

ビット10: 予約

ビット11: セットされている場合、イメージはWindows Driver Model (WDM) ドライバーである。

ビット12: 予約

ビット13: 画像がターミナルサーバーに対応している

AddressOfEntryPointは重要なものです。このメンバーには、イメージのエントリーポイント関数のRVAが含まれています (DLLにはエントリーポイントが必要ないので、これはNULLで構いません)。

それがすべてです。.text、.data、.bssなどの他のメンバーが何であるかに興味があるかもしれません。他にも、厄介な見た目の

まだ見ていないDataDirectoryのメンバー。

これらのメンバーについては後ほど詳しく見ていきます。今は、プログラムの実行を見てみましょう。

プログラムの実行

この段階では、プログラムを実行するだけであれば、すべての情報が提供されています。プログラムをロードした後、ローダがすべきことは、オプショナルヘッダからAddressOfEntryPointメンバを探し出し、そのアドレスを呼び出すことです。これはRVAであることを覚えておいてください。つまり、ローダはエントリーポイント関数へのリニアアドレスを得るために、このアドレスをImageBaseに追加する必要があります。

ここではその一例をご紹介します。

PEの実行ファイルの実行に必要なのは、これだけです :)

Data Directories

アブストラクト

リソース、シンボルテーブル、デバッグ情報、インポート、エクスポートテーブルなどは、オプションのヘッダにあるDataDirectoryメンバからアクセスできます。このメンバは、IMAGE_DATA_DIRECTORYの配列で、これらの情報を含む他の構造体にアクセスするために使用できます。

IMAGE_DATA_DIRECTORYの形式があります。

```
typedef struct _IMAGE_DATA_DIRECTORY {
    DWORD VirtualAddress;           // RVA of table
    DWORD Size;                    // size of table
} IMAGE_DATA_DIRECTORY, *PIMAGE_DATA_DIRECTORY;
```

DataDirectoryは、IMAGE_DATA_DIRECTORYの配列であることを覚えておいてください。この配列の各エントリでは、アクセスしたいさまざまなデータにアクセスする

ことができます。

インデックスのエントリーは以下の通りです。

0: 輸出用ディレクトリ

1: インポートディレクトリ

2: リソースディレクトリ

3: 例外ディレクトリ

4: セキュリティディレクトリ

5: ベースリロケーションテーブル 6: デバッグディレクトリ

7: 説明文字列

8: マシン値 (MIPS GP) 9: TLS
ディレクトリ

10: Load configuration directory 14:

COM+ data directory

例えば、エクスポートテーブルを読み取る必要がある場合は、DataDirectory[0]を参照します。リソースを読みたい場合は

DataDirectory[2].VirtualAddress:

それぞれのセクションには、特定のデータを解析するために必要な独自の構造があります。ここでは、その中でも特に便利なものを紹介します。

エクスポートテーブルの読み込み

エクスポートテーブルには、ライブラリやDLLからエクスポートされたすべての関数が含まれており、そのDLL内の関数アドレス、名前、序列番号などが記載されています。Win32 API関数のGetProcAddress()は、モジュールのエクスポートテーブルを序列番号または名前で解析し、そこからアドレスを返すことで動作します。このように、エクスポートテーブルを読むことは、便利な方法のひとつです。

エクスポート・テーブルを解析するには、まずエクスポート・ディレクトリ構造を取得する必要があります。これは、DataDirectory[0]を取得することで行います。

IMAGE_DATA_DIRECTORY構造体のVirtualAddressはRVAなので、イメージベースに追加する必要があることを覚えておいてください。次に exportDirectory

のポイントは、この素晴らしい構造にあります。

これは簡単なことです。AddressOfFunctionsは、関数アドレスの配列を指すRVAです。ただし、関数のアドレスもRVAです。AddressOfNamesは、関数名のリストへのポインタです。しかし、これらのアドレスはすべてRVAなので、関数名とアドレスを正しく取得するためには、イメージベースに追加する必要があります。

AddressOfNameOrdinalは、**序列のリストに対するRVA**です。**ordinals**は、エクスポートされた機能を表す单なる数字であり、アドレスではないため、RVAではありません。

エクスポートテーブルを正しく解析するには、ループで行う必要があります。例えば、以下のようにになります。

```
DWORD FunctionNameAddressArray = ((DWORD)ExportDirectory->AddressOfNames) + ((PBYTE)imageBase);
WORD FunctionOrdinalAddressArray = (WORD)ExportDirectory->AddressOfNameOrdinal + (PBYTE)imageBase;
PDWORD FunctionAddressArray = (PDWORD)ExportDirectory->AddressOfFunctions + (PBYTE)imageBase;

//! search for function in exports table
for ( i = 0; i < ExportDirectory->NumberOfFunctions; i++ )
{
    LPSTR FunctionName = FunctionNameAddressArray [i] + (PBYTE)imageBase;

    if (strcmp (FunctionName, funct) == 0) {

        WORD Ordinal = FunctionOrdinalAddressArray [i];
        DWORD FunctionAddress = FunctionAddressArray [Ordinal];
        return (PBYTE) (FunctionAddress + (PBYTE)imageBase);
    }
}
```

これを用いてGetProcAddress()を実装し、DLLのサポートに役立てることができます。

インポートテーブルの読み込み

輸出の表を読むだけでは不十分だったということですか？輸入の表を読むのはそれほど難しくありませんが、輸出の表よりも少し複雑です。OK、OK、輸入表を読むことに何の意味があるの？それは、読むことではなく、書くことです。プログラムのインポートテーブルにエントリを書き込むことで、GetProcAddress()を呼び出すことなく、ライブラリやDLL間での関数呼び出しが可能になります。Windowsでは、遅延ロードされたDLLやシステムDLLでこれを行います。

インポートテーブルを読むためには、インポートディレクトリ構造を見つける必要があります。これはDataDirectory[1]にあります。

```
PIMAGE_DATA_DIRECTORY DataDirectory = &OptionalHeader->DataDirectory [1];
PIMAGE_IMPORT_DESCRIPTOR importDirectory = (PIMAGE_IMPORT_DESCRIPTOR) (DataDirectory->VirtualAddress + ImageBase);
```

重要なのは、importDirectoryが記述子エントリの配列を指していることです。これらのエントリーはそれぞれ、インポートされたモジュール（インポートDLLなど）を表しています。では、この構造を見てみましょう。

```
    uint32_t Characteristics;           // 0 for terminating null import descriptor
    uint32_t OriginalFirstThunk;        // RVA to INT
};

    uint32_t TimeDateStamp;             // Time/Date of module, or other properties (see below)
    uint32_t ForwarderChain;           // Forwarder chain ID
    uint32_t Name;                   // Module name
    uint32_t FirstThunk;              // RVA to IAT (if bound this IAT has actual addresses)
```

ここで重要なのは、Name、OriginalFirstThunk、FirstThunkがRVAであることです。つまり、データを適切に解析するためには、イメージベースにアドレス（これらはポインタです）を追加する必要があります。Nameは、kernel32.dllのような、インポートされたモジュール名を指すRVAです。これはヌルで終端しています。

インポートディスクリプターの配列を扱っていることを覚えていますか？この配列にあるインポートディスクリプターの数を知るにはどうしたらいいでしょうか？配列はNULLのIMAGE_IMPORT_DESCRIPTORで終わっていますので、各エントリをループする簡単な方法は次のとおりです。

```
IMAGE_IMPORT_DESCRIPTOR* lpImportDesc;
while ( ! lpImportDesc->FirstThunk ) {

    //! work with lpImportDesc here
```

```

    lpImportDesc++; // move to next entry
}

TimeDateStampには、適切な時間/日付、または以下の値のいずれかを指定します。0:
モジュールがバインドされていない

```

-1: 画像が結合されている。実時間/日付スタンプの保存

ForwarderChainは、DLLの前方参照をサポートしている場合にのみ使用されます。これにより、DLL間の呼び出しを他のDLLに転送することができます。例えば、Windowsのkernel32.dllの一部の呼び出しは、他のDLLに転送されます。

FirstThunkはIATを指し、OriginalFirstThunkはインポートされたすべての関数を表す構造体の配列を指します。これは、Import Name Table (INT) です。これらのメンバーはいずれもRVAです。

そういうえば、もう一つの構造が出てきているのをご存知でしょうか。見てみましょう。

OriginalFirstThunkは、IMAGE_THUNK_DATA構造体の配列を指すRVAです。うわ、やった、また構造体だ。この構造体は小さいものですが。

```

typedef struct _IMAGE_IMPORT_BY_NAME
{
    uint16_t Hint;           // Possible ordinal number to use
    uint8_t Name[1];         // Name of function, null terminated
} IMAGE_IMPORT_BY_NAME, *PIMAGE_IMPORT_BY_NAME;

```

それがすべてです。第1パラメターは0でも構いませんが、これは関数がどのような序列番号を使用するかをローダに示唆するだけです。Nameは、関数の名前を表す文字の配列です。

ここからが本題です。IATは、機能を表すアドレスのリストに過ぎない。関数とは? このIMAGE_THUNK_DATA配列内の関数です。IMAGE_THUNK_DATA構造体を見てみると、関数名を表すユニオンがあるだけです。これは、インポートネームテーブル (INT) です。

例えば、IMAGE_THUNK_DATA[3]に格納されている関数の現在のアドレスを取得したいとします。そのアドレスは、IATの3番目のdwordで、IMAGE_IMPORT_DESCRIPTOR->FirstThunkで読み取ることができます。

そこで、機能名とアドレスを取得してみましょう。

```

unsigned int count=0;
while (lpThunk->u1.Function) {

    //! get the function name
    char* lpFunctionName = (char*)((uint8_t*)imageBase + (uint32_t)lpThunk->u1.AddressOfData.Name);

    //! go into the IAT to get this functions address
    uint32_t* addr = (uint32_t*)((uint8_t*)imageBase + lpImportDesc->FirstThunk) + count;

    // lpFunctionName now points to the null terminated function name
    // addr now points to the address of this function

    count++;
    lpThunk++;
}

```

イメージバインディング

ここからが面白いところです。IATには、ランタイム時にもビルド時にも、インポートされた関数のアドレスを入れることができます。束縛された画像とは、ビルド時にIATが関数に束縛される画像のことです。束縛されていない画像とは、ロード時にOSローダーによってIATが埋められた画像のことです。

画像に境界がある場合は、以下のようにして外部DLLの関数を呼び出すことができます。

```

__declspec (dllexport) void function ();
function () // calls myDll:function()

```

画像に境界がない場合、IATにはジャンクが含まれています。上記のコードを動作させるためにIATを更新するのは、OSローダーの責任です。これは、ロードされたDLLモジュールのエクスポートテーブルを読み込んで (GetProcAddress()を呼び出し)、そのインポート関数のIATエントリを上書きすることで実行できます。IATの上書きは、上記の手順で行うことができます - 関数のIATエントリを取得したら、それを上書きするだけです。)

この方法は、DLLや他のモジュールにフックをインストールする際にも有効です。

サポートリソース

はじめに

Windowsカーネルが、ディスクから何も読み込まずに画像を表示したり、XML設定ファイルを操作したりできることを不思議に思ったことはありませんか？リソースを追加する作業をしていて、それをOSでサポートできないかと考えたことはありませんか？答えは、"Yes, of course!" です。

しかし、リソースの解析は、他のディレクトリタイプよりも少し複雑です。他のセクションと同様に、ベースとなる

オプションヘッダの DataDirectory メンバから取得できる IMAGE_RESOURCE_DIRECTORY 構造体。

```
PIMAGE_DATA_DIRECTORY DataDirectory = &OptionalHeader->DataDirectory [2];
PIMAGE_RESOURCE_DIRECTORY resourceDirectory = (PIMAGE_RESOURCE_DIRECTORY) (DataDirectory->VirtualAddress + ImageBase);
```

これらのセクションへのアクセス方法にバターンがあることに気付きましたか？ そうだ、新しい構造にしよう。

```
typedef struct _IMAGE_RESOURCE_DIRECTORY {
    uint32_t Characteristics;
    uint32_t TimeStamp;
    uint16_t MajorVersion;
    uint16_t MinorVersion;
    uint16_t NumberOfNamedEntries;
    uint16_t NumberOfIdEntries;
    IMAGE_RESOURCE_DIRECTORY_ENTRY DirectoryEntries[1];
} IMAGE_RESOURCE_DIRECTORY, *PIMAGE_RESOURCE_DIRECTORY;
```

この構造では、最後の3つを除いて、興味深いフィールドはありません。

Win32リソースを扱ったことのある方は、リソースがIDや名前で識別できることをご存知かもしれません。この構造体の2つのメンバーは、これらのエントリの数と、エントリの総量 (NumberOfNamedEntries + NumberOfIdEntries) を知ることができ、これはすべてのエントリをループするのに便利です。お察しのとおり、エントリーはDirectoryEntriesの配列に入っています。DirectoryEntriesは、IMAGE_RESOURCE_DIRECTORY_ENTRY構造体の配列で構成されていて、以下のようなフォーマットになっています。

```
typedef struct _IMAGE_RESOURCE_DIRECTORY_ENTRY
{
    union {
        struct {
            DWORD NameOffset:31;
            DWORD NameIsString:1;
        };
        DWORD Name;
        WORD Id;
    };
    union {
        DWORD OffsetToData;
        struct {
            DWORD OffsetToDirectory:31;
            DWORD DataIsDirectory:1;
        };
    };
} IMAGE_RESOURCE_DIRECTORY_ENTRY, *PIMAGE_RESOURCE_DIRECTORY_ENTRY;
```

さてさて、これは醜い構造です。この構造は1つのリソース、つまりリソースディレクトリを表しています。

リソースディレクトリ構造

リソースかリソースディレクトリか？ちょっと立ち止まって考えてみましょう。（いいですか、リソースはツリーとして保存されていることを知っておくことが重要です。このツリーは次のような構造になっています。

ルートディレクトリ

```
リソースグループ1 ディレクトリ リソース1  
リソース2  
リソースグループ2 ディレクトリ リソース1  
リソース2  
リソースグループ3 ディレクトリ リソース1  
リソース2
```

...etc...

2 リソースグループにはいくつかの種類があり、そのグループに入っているリソースの種類を知ることができます。グループIDは以下の通りです。 1 - カーソル

- 3 - Bitmap
- 4 - Icon
- 5 - Menu
- 6 - Dialog
- 7 - String
- 8 - Fontdirectory
- 8 - Font
- 9 - Accelerator
- 10 - RcData
- 11 - Message table
- 16 - Version
- 17 - DlgInclude/li> 19
- Plug and Play 20 -

VXD

21 - アニメーションカーソル 22 - アニメーションアイコン 23 - HTML

24 - マニフェスト

リソースを見つけるためには、このツリーをトラバースする必要があります。ツリーには3つの層しかないと考えれば、難しいことではないというのが良い点です。まず、リソースディレクトリ内のすべてのエントリをループする方法を見てみましょう。

```
///! get first entry in directory
IMAGE_RESOURCE_DIRECTORY_ENTRY* lpResourceEntry = lpResourceDir->DirectoryEntries;

///! loop through all entries
int entries = lpResourceDir->NumberOfIdEntries + lpResourceDir->NumberOfNamedEntries;
while (entries-- != 0) {
```

```

//! look for bitmap resource (id=2)
if (lpResourceEntry->Id == 2) {
    //! see below
}
lpResourceEntry++;
}

```

これは簡単ですよね。IMAGE_RESOURCE_DIRECTORY_ENTRYのIdメンバーは、グループIDを格納するために使用されます。ビットマップを探すとしたら、ルートディレクトリのビットマップグループにあるはずなので、ID=2のエントリを探します。
IMAGE_RESOURCE_DIRECTORY_ENTRYは、リソースエントリとディレクトリの両方を表しているので、それが何であるかをどうやって見分けるのでしょうか？なぜかというと

もちろん、**DataIsDirectory**メンバーです。このメンバーが設定されていれば、それはディレクトリです。あ、でもディレクトリだったらどうやって読むの？ちょっと見てみましょう。

```

if (lpResourceEntry->DataIsDirectory) {
    lpResourceEntry = lpResourceEntry->OffsetToDirectory;
    lpResourceEntry += startOfResourceSection;
}

```

これは悪くありません。エントリがディレクトリの場合、上記は OffsetToDirectory から新しいディレクトリへのオフセットを取得し、それを startOfResourceSection に追加します。そうです…これはオフセットであって、RVAではないのです。そうなんですよね・・・。なぜ、マイクロソフト、なぜ！？

リソースセクションの開始点は、実際には IMAGE_RESOURCE_DIRECTORY_ENTRY 配列の最初のメンバーのアドレスになります。つまり、このアドレスを OffsetToDirectory で得たオフセットに加えれば、このディレクトリのIMAGE_RESOURCE_DIRECTORY構造体へのポインタを得ることができます。これで、ディレクトリのエントリを読み取る作業が開始されます。)

もし、特定のリソースのためにディレクトリを解析している最中であれば、ディレクトリ内のすべてのリソースエントリをループしてください。resourceEntry ID フィールドが、探そうとしているリソースID（ここではプログラム固有のID）と一致すれば、リソースデータが見つかったことになります。

リソースデータは、……ゾンデ！構造体に格納されています。ディレクトリエントリ構造のOffsetToData メンバから取得することができます。に似ています。

OffsetToDirectoryメンバーで、これもリソースセクションの開始からのオフセットです。

ポインタを取得すると、リソースデータを取り出すことができます。では、その構造を見てみましょう。

```

typedef struct _IMAGE_RESOURCE_DATA_ENTRY
{
    uint32_t OffsetToData;
    uint32_t Size;
    uint32_t CodePage;
    uint32_t Reserved;
} IMAGE_RESOURCE_DATA_ENTRY, *PIMAGE_RESOURCE_DATA_ENTRY;

```

その通りです。OffsetToDataは実際のリソースデータを指すRVAで、Sizeはそのデータのサイズをバイト単位で表しています。例えば、ビットマップのリソースを探している場合、OffsetToDataはビットマップのBITMAPINFOHEADER構造を指すRVAとなり、これは任意のビットマップローダで処理することができます。

Conclusion

この章はこれで終わりです。今後は、デバッグデータやCOMDATSなどの項目を追加していく予定です。

この章にはデモはありません。この章は主に、PE実行ファイル形式の内部動作とその作業に興味がある人のために公開されています。メインのシリーズでは、プログラムのロードと実行だけを行うことがありますので、その他の情報は完全性のためにのみ提供されています。デモのためにテキストで提供されたコードは、すべて動作確認済みです（若干の修正を加えています）。

次の章では、PE実行ファイルフォーマットの使用と、ユーザーモードプログラムをサポートするローダーの構築を行います。続いて、マルチタスクについて説明します。

次の機会まで。

~マイク

BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中 質問やコメント
は？お気軽にご連絡ください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。



第23章 ホーム

Chapter 23

Home



オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - 8259A PICマイクロコントローラ

by Mike, 2007

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

8259A PICマイクロコントローラの全ピンが表示されています。

Introduction

歓迎します。:)

このチュートリアルでは、非常に重要なトピックを取り上げています。プログラム可能なインタラプトコントローラです。このマイクロコントローラを初期化して、IRQにマッピングする必要があります。これは、割り込みを設定したり、割り込み要求を処理する際に必要となります。

今回は、初めてのコントローラチュートリアルです。このチュートリアルでは、それぞれのデバイスについて深く掘り下げ、それらを扱うための実用的なインターフェイスを構築しています。保護モードに入っているため、ガイドとなるものがないことを忘れないでください。一步間違えると、予想外の結果になってしまいます。手がかりがない分、それぞれのコントローラーと直接コミュニケーションをとらなければなりません。そのため、本連載ではハードウェア・プログラミングの概念を重視し、読者がハードウェア・レベルのプログラミングをより深く理解できるようにしています。

このチュートリアルでは、学んだことをすべて試してみます。このチュートリアルでは、学習したことすべて試すことができます。できるだけ簡単にすることにします。いいですか？

Get Ready

これは、多くのマイクロコントローラ・プログラミング・チュートリアルの最初のものです。このチュートリアルでは、各マイクロコントローラのほぼすべての機能をカバーしています。メインシリーズでは、必要に応じてこれらのチュートリアルを参照し、これらのコントローラが必要とするものをカバーしていきます。

このチュートリアルはかなり複雑です。8259Aマイクロコントローラをハードウェアとソフトウェアの両方の観点から取り上げ、PCとの接続や動作を正確に理解します。また、このマイクロコントローラのすべてのコマンド、レジスタ、および部品をカバーします。

History

To do - 近いうちにこのセクションを追加する予定です。

8259A PICはハードウェア割り込みを扱うため、まずは割り込みとは何か、どのように動作するのかを基本的に理解する必要があります。

Interrupts

割り込みとは、ソフトウェアやハードウェアの注意を必要とする外部の非同期信号のことです。これにより、現在のタスクを中断して、より重要なことを実行することができます。

ハードにはならない。割り込みは、ゼロ除算などの問題をトラップするのに役立つ方法を提供します。プロセッサが現在実行中のコードに問題を発見した場合、その問題を解決するために実行する代替コードをプロセッサに提供します。

その他の割り込みは、ソフトウェアをルーチンとしてサービスする方法を提供するために使用することができます。これらの割り込みは、システム内の任意のソフトウェアから呼び出すことができます。これは、リング3のアプリケーションがリング0レベルのルーチンを実行する方法を提供するシステムAPIによく使われます。

特に、非同期に状態が変化する可能性のあるハードウェアから情報を受け取る方法として、インタラプトは多くの用途があります。

割り込みの種類

割り込みには、「ソフトウェア割り込み」と「ハードウェア割り込み」の2種類があります。

ソフトウェアインタラプト

ソフトウェア割り込みとは、ソフトウェアで実装・起動される割り込みのことです。通常、プロセッサの命令セットには、ソフトウェア割り込みを処理するための命令が用意されています。x86アーキテクチャの場合、これらは通常INT imm、INT 3です。また、IRET、IRETD命令も使用します。

例えば、ここではソフトウェア命令で割り込みを発生させます。

これらの命令は、ソフトウェアによる割り込みの生成や、割り込みルーチン (IR) の実行に使用できます。

ここでは、ソフトウェア割り込みは取り上げません。8259A PICマイクロコントローラは、ハードウェア割り込みのみ対応しています。ソフトウェア割り込みについては、別のチュートリアルで取り上げます。

ハードウェアインタラプト

ハードウェア割り込みとは、ハードウェアデバイスによって引き起こされる割り込みのことです。通常、これらは注意を必要とするハードウェアデバイスです。ハードウェアインターブトハンドラは、このハードウェア要求を処理するために必要となります。

スプリアスインターラプト

これは、割り込みラインの電気的干渉や、ハードウェアの不具合によって発生するハードウェア割り込みです。これは絶対に避けたいことです。

割り込みモード

割り込みにはいくつかのモードとクラスがあり、それらをカバーする必要があります。PICのプログラミングでは、モードを選択する必要があります。
注：このセクションでは、8259A PICのハードウェア・ピン・レイアウトの知識が必要になるかもしれません。これについては次のセクションで説明します。レベルトリガ

レベルトリガ割り込みは、PICのIR（Interrupt Request）ラインに電流（1）が流れたときに発生すると判断されます。あるデバイスが信号を送る（この設定ラインをアクティブにして）、割り込みが処理されるまでその状態を維持します。

レベルトリガの割り込みラインは、回路がそれに対応できるように設計されていれば、複数の割り込みで共有することができます。

このモードは、ラインを共有する方法のため、好ましいモードです。IRラインがアクティブになると、CPUは同じラインを共有しているすべてのデバイスを検索し、どのデバイスが信号をアクティブにしているかを見つけ出します。見つかった後、CPUはすべての機器を再チェックし、サービスを必要とする他の機器がないことを確認します。

この方法の問題点は、より高い優先順位で処理しなければならない割り込みがあった場合、他の割り込みが処理されるまで、他のすべての割り込みが永久にブロックされてしまうことです。結局のところ、一度にアクティブにできるのは1つのラインだけです。

エッジトリガ

エッジトリガ割り込みは、PICのIR（Interrupt Request）ラインに電流（1）が流れたときに発生するとされている。デバイスはシングルパルスで信号を送り（このラインをアクティブにする）、ラインを元の状態に戻します。

エッジトリガの割り込みラインは、回路がそれに対応できるように設計されていれば、複数の割り込みで共有することができます。パルスが短すぎて検出できない場合は、検出されません。

これは、割り込み要求を通知する電流パルスに過ぎないため、エッジトリガモードでは、レベルトリガのようにIRQラインを共有する問題は発生しません。

もちろん、IRQラインには1パルスの電流が流れるだけなので、割り込みを見逃す可能性もあります。これは、初期のコンピュータでCPUのロックを引き起こす原因となりました。しかし、最近ではこのようなロックアップは時間の経過とともに減少しています。

ハイブリッド

この2つのモードにはそれぞれ長所と短所があります。多くのシステムでは、この2つのモードのハイブリッドを採用しています。具体的には、ほとんどのシステムでは、CPUのNMI（Non Maskable Interrupt）ピンでエッジトリガとレベルトリガの両方の割り込みをチェックしています。この目的は、NMIピンがシステムの重大な問題を知らせるために使用され、大きな問題やシステム全体の誤動作、場合によってはハードウェアの損傷を引き起こす可能性があるからです。

Non Maskable Interrupt（ノンマスクブル・インターブト）とは、どのデバイスからも無効にされたり、マスクされたりすることのない割り込みのことです。これにより、ハイブリッドセットアップと同様に、NMIピンがセットされた場合、システムは大きな問題を起こすことなく穏やかに終了することができます。

シグナルされたメッセージ

この種のハードウェア割り込みは、物理的な割り込みラインを使用しません。このタイプのハードウェア割り込みは、物理的な割り込みラインを使用せず、システムバスなどの別の媒体を利用してメッセージを送信します。このタイプの割り込みは、エッジトリガ割り込みと同様に、デバイスが媒体上にパルス状の電流を送るだけです。

この種のシステムでは、制御バス上にメッセージ信号による割り込み番号を示す特別な割り込みラインを使用することができます。これらの番号は、一連のビットとしてメディア上で送信されるため、他の割り込みタイプのように1本の割り込みラインに制限されることはありません。また、このタイプの割り込みは、他の割り込みタイプのように、1本の割り込みラインに限定された制約がないため、下層システムが許す限り、多くの割り込みを管理することができます。また、これらのタイプの割り込みは、割り込みベクターの共有にも対応しています。

PCI Expressでは、これらのタイプの割り込みを使用しています。

以上です。

さて、ここには多くの情報があります。）8259Aは、レベルトリガとエッジトリガの割り込みしかサポートしていません。このため、8259Aマイクロコントローラを扱う際には、これに主眼を置く必要があります。

Interrupt Vector Table

IVT（Interrupt Vector Table）は、インターブトベクターのリストです。IVTには256個のインターブトがあります。

割り込みルーチン(IR)

割り込みルーチン(IR)は、割り込み要求(IRQ)を処理するための特別な機能です。

プロセッサがINTなどの割り込み命令を実行すると、IVT（Interrupt Vector Table）内のその位置でIR（Interrupt Routine）が実行されます。

つまり、私たちが定義したルーチンを実行するだけです。難しくないでしょう？この特別なルーチンは、AXレジスタの値に基づいて、通常実行する割込み関数を決定します。これにより、1つのインターブトコールに複数のファンクションを定義することができます。例えば、DOSのINT21h関数0x4c00のように。

覚えておいてください。割り込みを実行すると、自分が作成した割り込みルーチンが単純に実行されます。例えば、INT 2という命令は、IVTのインデックス2のIRを実行します。いいですか？

IVTマップ

IVTは、物理メモリの最初の1024バイト、アドレス0x0から0x3FFまでに配置されています。IVT内の各エントリは4バイトで、次のような形式になっています。

バイト0: 割り込みルーチン(IR)のオフセットロードレス バイト1:

IRのオフセットハイアドレス

バイト2 : IRのセグメントLowアドレス

バイト3 : IRのセグメントハイアドレス

IVTの各エントリには、単に呼び出すIRのアドレスが含まれていることに注意してください。これにより、メモリ上の任意の場所 (Our IR) に簡単な関数を作成することができます。

IVTに関数のアドレスが格納されていれば、すべてがうまくいきます。

では、IVTを見てみましょう。最初の数個の割り込みは予約されており、そのままです。

x86 Interrupt Vector Table (IVT)		
Base Address	Interrupt Number	Description
0x000	0	Divide by 0
0x004	1	Single step (Debugger)
0x008	2	Non Maskable Interrupt (NMI) Pin
0x00C	3	Breakpoint (Debugger)
0x010	4	Overflow
0x014	5	Bounds check
0x018	6	Undefined Operation Code (OPCode) instruction
0x01C	7	No coprocessor
0x020	8	Double Fault
0x024	9	Coprocessor Segment Overrun
0x028	10	Invalid Task State Segment (TSS)
0x02C	11	Segment Not Present
0x030	12	Stack Segment Overrun
0x034	13	General Protection Fault (GPF)
0x038	14	Page Fault
0x03C	15	Unassigned
0x040	16	Coprocessor error
0x044	17	Alignment Check (486+ Only)
0x048	18	Machine Check (Pentium/586+ Only)
0x05C	19-31	Reserved exceptions
0x068 - 0x3FF	32-255	Interrupts free for software use

難しいことではありません。これらの割り込みは、それぞれIVT内のベースアドレスに配置されています。

プロジェクトモード(PMode)での割り込み処理

プロジェクトモードでは、各IVTエントリが、IDT (Interrupt Descriptor Table) 内で定義された割り込みルーチン (IR) を指す必要があります。IDTについては、このチュートリアルとは直接関係がないため、別のチュートリアルで詳しく説明します。

IDTは、実行する割り込みルーチン(IR)のベースアドレスを記述した割り込み記述子の配列で、保護レベルやセグメント情報などの追加情報を含んでいます。PModeでは、使用されるメモリマップを定義するグローバルディスクリプターテーブル (GDT) を使用します。割り込みルーチンの多くは、GDTによってマップされたコード記述子の中に入ります。これがPModeでIDTが必要な理由です。

今はまだ理解できなくても気にしないでください。とりあえず、256個の関数ポインタの配列で、IVTと同じようにマッピングされていると思ってください（普通はそうです）。

Hardware Interrupts

割り込みには、ソフトウェアで発生させるもの (INT、INT 3、BOUND、INTOなどの命令で使用) と、ハードウェアで発生させるものがあります。

ハードウェア割り込みは、PCにとって非常に重要です。これにより、他のハードウェアデバイスが、何かが起ころうとしていることをCPUに知らせることができます。例えば、キーボードのキーストロークや、内部タイマーの1クロック分の目盛りなどです。

これらの割り込みが発生したときに、どのようなIRQ (Interrupt Request) を生成するかをマッピングする必要があります。こうすることで、ハードウェアの変化を追跡することができます。では、これらのハードウェア割り込みについて見てみましょう。		
x86 Hardware Interrupts		
8259A Input pin	Interrupt Number	Description
IRQ0	0x08	Timer
IRQ1	0x09	Keyboard
IRQ2	0x0A	Cascade for 8259A Slave controller
IRQ3	0x0B	Serial port 2
IRQ4	0x0C	Serial port 1
IRQ5	0x0D	AT systems: Parallel Port 2. PS/2 systems: reserved
IRQ6	0x0E	Diskette drive
IRQ7	0x0F	Parallel Port 1
IRQ8/IRQ0	0x70	CMOS Real time clock
IRQ9/IRQ1	0x71	CGA vertical retrace
IRQ10/IRQ2	0x72	Reserved
IRQ11/IRQ3	0x73	Reserved
IRQ12/IRQ4	0x74	AT systems: reserved. PS/2: auxiliary device
IRQ13/IRQ5	0x75	FPU
IRQ14/IRQ6	0x76	Hard disk controller
IRQ15/IRQ7	0x77	Reserved

各デバイスについては、まだあまり気にする必要はありません。8259Aのピンについては、次のセクションで詳しく説明します。この表に記載されている割り込み番号は、これらのイベントが発生したときに実行されるデフォルトのDOS割り込み要求 (IRQ) です。

ほとんどの場合、新しい割り込みテーブルを作り直す必要があります。そのため、ほとんどのオペレーティングシステムでは、PICが使用する割り込みをリマップして、IVT内の適切なIRQを呼び出すようにする必要があります。これは、リアルモードのIVTではBIOSが行ってくれます。このチュートリアルでは、後ほどこの方法を説明します。

8259 Programmable Interrupt Controller

8259マイクロコントローラ・ファミリは、プログラマブル・インターラブト・コントローラ(PIC)集積回路(IC)のセットです。チュートリアル7をもう一度見てみましょう… の下に

「プロセッサ・アーキテクチャ」のセクションでは、プロセッサが独自のPICマイクロコントローラを内蔵していることに注目してください。これは非常に重要な点です。

回路設計上の制約から、PICは8つのIRQしかサポートしていません。これは大きな制限である。デバイスが増えるにつれ、IBMはこの制限が非常に悪いものであることにすぐに気づきました。そのため、ほとんどのマザーボードにはセカンダリ (スレーブ) PICマイクロコントローラが搭載されており、プライマリPICと一緒に動作します。

プロセッサを使用しています。今日では、これは非常に一般的なことです。1つのPICは、別のPICと「カスケード接続」（一緒に動作可能）することができます。これにより、追加のPICで多くのIRQをサポートすることができます。

対応するPICの数が多いほど、より多くのIRQを処理できます。カスケード接続して最大64のIRQをサポートすることができます。いいですね。

覚えておいてください。ほとんどのコンピュータには2つのPICがあり、1つはプロセッサ内部に、1つはマザーボードに搭載されています。システムによってはこれがない場合もあります。覚えておいてください。各PICは最大8つのIRQをサポートしています。

覚えておいてください。各PICは相互に通信することができ、PICの数に応じて最大64のIRQが可能です。

難しいことではありません :)

8259 Hardware

マイクロコントローラがハードウェアレベルでどのように動作するかを理解することは、ソフトウェア側の動作を理解するのに役立ちます。PICはハードウェアの割り込み時にのみ使用されることを覚えておいてください。

8259Aマイクロコントローラ

このチュートリアルの上部には、すべての電子ピンが表示された実際の8259 Dual Inline Package (DIP)の画像があります。ここでは、より分かりやすくするために、コントローラをよりシンプルなグラフィックで表現します。これらの図には表示されていませんが、8259が持つ唯一のピンは、GND (グランド) とVcc (入力電圧) です。これらのピンは、このチュートリアルの一番上の写真に表示されています。

まず、これからプログラミングする内容について説明します。

8259Aプログラマブルインタラプトコントローラ』です。

上の画像の各線は、コントローラの各電子ピンを表しています。これらの電子ピンは、コントローラと他のシステムとの接続に使用されます。

このチップは、OS内のIRQを処理するためにプログラムする必要があります。各ピンで詳しく見てみましょう。重要なピンは太字にしました。

WR端子：ライトストロープ信号に接続します (Pentiumでは8本のうちの1本) RD端

子：IOCR (Input Output Control Routine) 信号に接続します。INT端子：マイクロプロセッサのINTR端子に接続します。

INTA端子：マイクロプロセッサーのINTA端子に接続します。AO端子：

異なるコマンドワードを選択可能

CS端子：プログラミングや制御のためにチップを有効にする。

SP/EN端子：スレーブプログラム (SP) / イネーブルバッファ (EN)。

スレーブプログラム (1=マスター、0=スレーブ)

Enable Buffer (バッファモード時のデータバストランジを制御)

CAS0,CAS1,CAS2端子です。カスケード接続されたシステムにおいて、マスターのPICコントローラーからスレーブのPIC

コントローラへの出力に使用します。D0～D7端子。D0～D7端子：8ビットのデータコネクタ端子です。

ここには、いくつかの重要なピンがあります。ピンD0～D7は、外部デバイスがPICと通信するための手段です。これは小さなデータバスのようなもので、PICにデータを送信する方法を提供します。

PIC同士を接続できることを覚えておいてください。これにより、最大64個のIR番号をサポートすることができます。言い換れば、64個のハードウェア割り込みです。CAS0、CAS1、CAS2ピンは、これらのPICの間で信号を送信する方法です。

INTピンとINTAピンを見てください。プロセッサのINTピンとINTAピンがPICのこれらのピンに接続されていることを「プロセッサの視点」の項で思い出してください。割り込みを実行しようとするとき、プロセッサはFLAGSレジスタの割り込みフラグ (IF) とトランプフラグ (TF) をクリアして、INTRピンを無効にすることを覚えておいてください。PICのINTピンは、プロセッサのINTRピンに接続します。

つまり、プロセッサは、割り込みを実行する際に、基本的にPICのINTピンを無効にします。

これにより、ピンIR0-IR7を他のPICにストリームすることができます。これらの8つのピンは、実行される8ビットの割り込み番号を表しています。これらのラインは、他のPICコントローラに割り込み番号を送信する方法を提供し、そのコントローラが代わりに処理できるようにします。

ここで重要なのは、複数のPICを組み合わせることで、より多くの割り込みルーチン数をサポートできるということです。IRラインは、他のPICのデータラインに接続して、データを転送します。ラインは8本 (8ビット) しかないので、最大で8台のPICを接続し、64個の割り込み番号をサポートすることができます。

なるほど...。ここにはたくさんのが書かれていますね。プロセッサがプライマリPICに接続する方法、PICが他のPICと結合してPICのチャイを作成する方法を説明しました。

これは素晴らしいことです、全く役に立ちません。割り込みはどのようにしてハードウェアを介して実行されるのか? このコントローラが「プログラマブル」なのはなぜか? どうやってPICをプログラムして、自分のニーズに合わせて働かせることができるのか?

PICのプログラミングは、PICが持つ8ビットのデータラインを通じてコマンドバイトを送信することを中心に行われます。この8ビットのコマンドバイトは、PICが何をすべきかを示す特定のフォーマットに従っています。PICをプログラムするためには、これらのコマンドを知る必要があります。これについては後ほど説明します。

PICがどのように動作するかを詳しく見てみましょう。これは、8259Aのピンや、割り込み信号がどのように送られるかを理解するのに役立ちます。

8259A コネクション

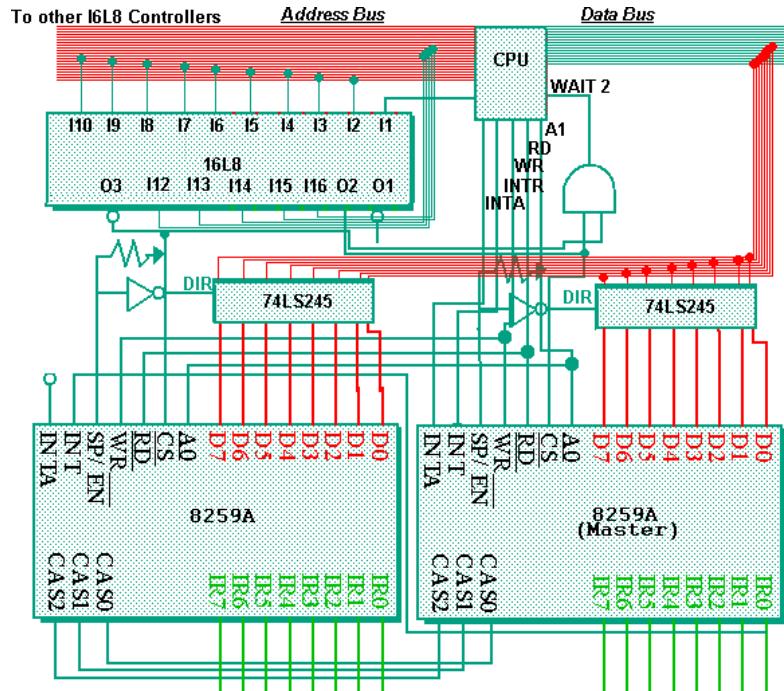
注：このセクションでは、デジタル・ロジック・エレクトロニクスの知識が必要な場合があります。

なるほど...。ここまででは、8259AのPICピンを見てきました。では、これらのピンを別の視点から見て、一般的なコンピュータの中でどのように見えるかを見てみましょう。

PICとプロセッサーの接続

まず、現在のほとんどのコンピュータには8259A PICが2台搭載されていると言ったことを覚えていますか？これは半分しか当てはまりません。Processor Architectureで、プライマリPICがプロセッサに組み込まれていることを思い出してください。これには理由があって、すぐにわかります。

簡単に説明すると、今回のシステムには2つのPICコントローラがあり、どちらもマザーボード上に直接設置されているとします（いずれもプロセッサとは統合されていません）。これをグラフにしてみると、こんな感じになります。



なるほど....ここには多くのことが起こっています。これはIOサブシステムとISAバスの一部を表示しており、8259Aコントローラが共通の16L8を介してどのようにシステムバスに接続するかを示しています。

このシリーズでは、デジタル・ロジック・エレクトロニクスは必須条件ではありませんので、分からなくても全く気にしないでください（笑）。また、この画像は細部が不足しています。とはいっても、部品間の基本的なリンクや接続は表示されています。

上の画像を見ると、いくつかの重要な注意点があります。

スレーブコントローラがプライマリコントローラにどのように接続しているかに注目してください。

プロセッサと直接接続する必要があるのはプライマリPICだけであることに注目してください。最近のコンピュータでは、このような緊密な統合が必要なため、プライマリPICをプロセッサの内部に直接組み込んで、この依存関係を最小限にしています。

また、CAS0-CAS2ピンが2番目のPICに直接接続されていることにも注目してください。これにより、プライマリPICがセカンダリPICにコマンドを送ることができます。

そして、周知の通り、IRラインはそのラインを制御する他のコントローラに接続されています。例えば、ハードウェア割り込み0は、タイマーの割り込みを表します。8254プログラマブルインターバルタイマ（PIT）コントローラは、直接接続されているため、IR0ラインを通してプライマリPICに信号を送ります。この信号は、割り込みが処理されるまでアクティブな状態を維持する電流か、一定時間保持される単一パルスのいずれかになります。PICに何を監視させたいのかを制御することができます。これについては、後で詳しく説明します。さて、ここからが本題です。） 悪名高き8259A PICマイクロコントローラです。

ハードウェア割り込みの実行方法

すべてのマイクロプロセッサーの下側には、コネクターがあります。これは平板状のものと、マザーボードに接続するためのピン状のものがあります。これらのピンのうち2つは、INTRとNMIのピンです。これに加えて、割り込みの完了を確認するためのもう1つのピン、INTAがある。

ソフトウェア割り込みは、ハードウェア割り込みとは扱いが異なります。どちらの割り込みも、メモリ上のアドレス0～0x3ffにある「割り込みベクターテーブル」の中に入っています。覚えておいてください。プログラマブルインターバルコントローラで処理されるのは、ハードウェアインタラプトのみです。

割り込みの発生

- デバイスコントローラが割り込みを発生させる必要がある場合、何らかの方法でPICに信号を送る必要があります。ここでは、このデバイスをタイマーとし、割り込みライン0を使用すると仮定します。

2. The timer controller signals the PIC by activating the IR0 line. This changes its state from a 0 (No power) to a 1 (Power is going through the line.)
3. The PIC sets the bit representing the IRQ inside of the **Interrupt Request Register (IRR)**. In this example, bit 0 will be set to 1.
4. The PIC examines the **Interrupt Mask Register (IMR)** to see if the interrupt can be serviced.

割り込みの処理が可能な場合、PICは処理を待っている優先度の高い割り込みがあるかどうかを判断します。優先度の高い割り込みがある場合は、優先度の高い割り込みが処理されるまで割り込み要求は無視されます。

5. 割り込みの処理が可能で、より優先度の高い割り込みがない場合、PICは次のステップに進みます。

6. The PIC signals the processor through the INTA pin to inform the processor an interrupt has been fired.

The processor now knows that an interrupt has been fired.

1. プロセッサは、割り込みを認識します。

2. The CPU completes execution of the current instruction.

3. The CPU examines the **Interrupt Flag (IF)** within **RFLAGS**.

4. IFがセットされている場合、CPUはINTRピンを介して割り込み要求をPICに返します。IFがクリアされた場合、割り込み要求は無視されます。

5. The PIC receives the acknowledgment signal through INTR.

6. The PIC places the interrupt vector number into the D0-D7 pins.

7. この割込みベクター番号は、PICの初期化時にICW (Initialization Control Word) 2から取得します。これについては後で説明します。

8. The PIC also places the IRQ number into D0-D7

9. The PIC sets the correct bit inside of the **In Service Register (ISR)**. In this case, it is bit 0. This indicates that Interrupt 0 is currently being serviced.

これで、プロセッサは実行すべきIRQ番号と割り込みベクター番号を手に入れたことになります。

1. インタラプション

2. The processor interrupts the current process. It pushes EFLAGS, CS, and EIP on the stack.
3. The processor uses the interrupt vector number (given by the PIC).

リアルモードでは、CPUはIVTにオフセットします。プロテクトモードでは、プロセッサはIDTにオフセットされます。リアルモードでは

CPUがオフセットしてIVTに正しく入力されていること

CPUは呼び出したい割込みのベースアドレスをCS:IPにロードする 割込み
が制御を行う。

プロテクトモード。

CPUは、ロードされたIDTを使って、オフセットして

ゲートディスクリプタのセレクタ・ファイルは、CSセグメントセレクタに読み込まれます。

ゲートディスクリプタのoffset fieldはEIPに読み込まれます。

ページングが有効な場合、このアドレスはリニアアドレスから物理アドレスに変換されます。ここで、

CPUは現在の状態に対してアーキテクチャ固有のセキュリティチェックを行います。

ここで、割り込みルーチンは、ゲートディスクリプタ+CS:EIPから制御を受けることができます。

割り込みサービスルーチン

現在、ISRはハードウェア割り込みを処理するために実行されています。ISRは、特定のデバイスを処理するために必要なあらゆる動作を行うことができます。例えば、デバイスへのデータの読み書き、ステータスレジスタの読み出し、コマンドの送信などです。

この間、すべての割り込みはIMR (Interrupt Mask Register) によってマスクアウトされます。つまり、割り込みを終了させる要求が出されるまで、すべてのハードウェア割り込みが禁止されます。このためには、EOI (End of Interrupt) コマンドをPICに送信する必要があります。

EOI信号がプライマリPICのコマンドレジスタを介してPICに送信された後、PICはインサービスレジスタ (IRR) の appropriateビットをクリアし、新たな割り込みを処理する準備が整います。

その後、割り込みサービスルーチンはIRETD命令を実行し、割り込みが発生したときにプロセッサがプッシュしたEFLAGS、CS、EIPレジスタをポップします。これにより、第一のタスクに制御が戻ります。

8259A Registers

8259Aは、プロセッサと同様に、いくつかの内部レジスタを持っています。

コマンドレジスタ

これは、マイクロコントローラーにコマンドを送信するための書き込み専用のレジスタです。送信可能なコマンドには様々なものがあります。いくつかのコマンドは、他のレジスタからの読み出しに使用され、他のコマンドは、初期化やデータの送信に使用されます。ここでは、これらのコマンドについて説明します。

ステータスレジスタ

これは、PICのステータスを決定するためにアクセスすることができるリードオンリーのレジスタです。

割り込み要求レジスタ (IRR)

このレジスタは、確認応答を保留している割り込みを指定します。

注：このレジスタは内部にあり、直接アクセスすることはできません。

Interrupt Request Register (IRR)		
Bit Number	IRQ Number (Primary controller)	IRQ Number (Slave controller)
0	IRQ0	IRQ8
1	IRQ1	IRQ9
2	IRQ2	IRQ10
3	IRQ3	IRQ11
4	IRQ4	IRQ12
5	IRQ5	IRQ13
6	IRQ6	IRQ14
7	IRQ7	IRQ15

ビットがセットされている場合は、デバイスからの割り込み信号があり、PICはCPUに信号を送りましたが、割り込みを進めるためにCPUからの確認を待っている状態です。

イン・デバイス・レジスタ (ISR)

このレジスタは、すでにアクノリッジされていて、EOI (End of Interrupt) 信号を待っている割り込みを指定します。EOI信号は、割り込みの終了を決定する非常に重要な信号です。

注：8259Aに割り込みを認識させるために、割り込み完了時にEOI信号を送る必要があります。そうしないと、未定義の動作や誤動作の原因になります。これについては後で説明します。

注：このレジスタは内部にあり、直接アクセスすることはできません。

In Service Register (ISR)		
Bit Number	IRQ Number (Primary controller)	IRQ Number (Slave controller)
0	IRQ0	IRQ8
1	IRQ1	IRQ9
2	IRQ2	IRQ10

3	IRQ3	IRQ11
4	IRQ4	IRQ12
5	IRQ5	IRQ13
6	IRQ6	IRQ14
7	IRQ7	IRQ15

割り込みマスクレジスタ(IMR)

これにより、このレジスタで指定された割り込みを実行する前に、特定のより重要な割り込みを実行することに集中することができます。

これは8ビットのレジスタで、各ビットが割り込みを無効にするかどうかを決定します。このビットが0の場合、割り込みは有効です。1であれば、割り込みデバイスは無効になります。

Interrupt Mask Register (IMR)		
Bit Number	IRQ Number (Primary controller)	IRQ Number (Slave controller)
0	IRQ0	IRQ8
1	IRQ1	IRQ9
2	IRQ2	IRQ10
3	IRQ3	IRQ11
4	IRQ4	IRQ12
5	IRQ5	IRQ13
6	IRQ6	IRQ14
7	IRQ7	IRQ15

これは重要なレジスタで、特定のデバイスからの割り込みを有効にしたり無効にしたりすることができます。これらのIRQはそれぞれ、上に示したx86 Hardware Interruptsの表に記載されているデバイスを表しています。

例えば、COM1（シリアルポート1）を有効にしたいとします。x86ハードウェア割り込みテーブルを見ると、これはIRQ4にマッピングされています。したがって、COM1の割り込みを有効にするためには、プライマリPICの割り込みマスクレジスタのIRQ4ビットを設定すればよいことになります。このレジスタはソフトウェアのポート番号0x21にマッピングされています（この点については後述します）ので、このポート位置に書き込むことでビットを設定するだけです。

```

in    al, 0x21          ; read in the primary PIC Interrupt Mask Register (IMR)
and   al, 0xEF          ; 0xEF => 11101111b. This sets the IRQ4 bit (Bit 5) in AL
out   0x21, al          ; write the value back into IMR

```

かっこよすぎて学校に行けないB)

ハードウェア割り込みが発生すると、8259AはEOI (End of Interrupt) 信号を受信するまで、他のすべての割り込みをマスクします。割り込みの完了時にEOIを送信する必要があります。これについては後述します。

8259A Software Port Mappings

他のハードウェアコントローラと同様に、BIOSのPOSTでは、各コントローラがソフトウェアポートの特定の領域を使用するようにマッピングされます。このため、PICコントローラと通信するためには、ソフトウェアポートを使用する必要があります。

8259A Software Port Map	
Port Address	Description
0x20	Primary PIC Command and Status Register
0x21	Primary PIC Interrupt Mask Register and Data Register
0xA0	Secondary (Slave) PIC Command and Status Register
0xA1	Secondary (Slave) PIC Interrupt Mask Register and Data Register

Primary PICのInterrupt Mask RegisterがPort 0x21にマッピングされていることに注目してください。どこかで見たことがあるような気がしますね。

コマンドレジスターとステータスレジスターは、同じポート番号を共有する別のレジスターです。コマンドレジスターは書き込み専用で、ステータスレジスターは読み出し専用です。これは重要な違いで、PICは書き込みラインと読み込みラインのどちらがセットされているかによって、アクセスするレジスタを決定します。

各デバイスのレジスタと通信したり、PICを制御したりするためには、これらのポートに書き込みができる必要があります。それでは、PICのコマンドを見てみましょう。

8259A Commands

PICの設定は非常に複雑です。PICの設定は、初期化や動作に使用される様々な状態を含むビットパターンである、一連のCommand Wordsを通して行われます。少し複雑に見えるかもしれません、それほど難しいことではありません。

そのため、まずはPICコントローラを初期化し、その後、PICを操作・制御する方法をご紹介します。

初期化コントロールワード (ICW)

PICを初期化する目的は、PICのIRQ番号を我々のものに再マッピングすることです。これにより、ハードウェア割り込みが発生したときに、適切なIRQが生成されるようになります。

PICを初期化するためには、プライマリPICコマンドレジスタにコマンドバイト (Initialization Control Word (ICW) として知られている) を送信する必要があります。これがICW 1です。

最大で4つの初期化制御ワードがあります。これらは必須ではありませんが、しばしば必要となります。ここでは、その内容をご紹介します。

注：システム内に複数のPICがあり、互いにカスケード接続される場合は、両方のPICにICWを送信する必要があります。

ICW 1

これは、PICを初期化するために使用される一次制御ワードです。これは、一次PICコマンドレジスタに入れなければならない7ビットの値です。これは、フォーマットです。

Initialization Control Word (ICW) 1		
Bit Number	Value	Description
0	IC4	If set(1), the PIC expects to receive IC4 during initialization.
1	SNGL	If set(1), only one PIC in system. If cleared, PIC is cascaded with slave PICs, and ICW3 must be sent to controller.
2	ADI	If set (1), CALL address interval is 4, else 8. This is usually ignored by x86, and is default to 0
3	LTIM	If set (1), Operate in Level Triggered Mode. If Not set (0), Operate in Edge Triggered Mode
4	1	Initialization bit. Set 1 if PIC is to be initialized

5	0	MCS-80/85: Interrupt Vector Address. x86 Architecture: Must be 0
6	0	MCS-80/85: Interrupt Vector Address. x86 Architecture: Must be 0
7	0	MCS-80/85: Interrupt Vector Address. x86 Architecture: Must be 0

ご覧のように、ここにはたくさんのが起こっています。これらのうちのいくつかは以前に見たことがあります。これらのビットのほとんどはx86プラットフォームでは使用されていないので、これは思ったほど難しくありません。

プライマリPICを初期化するためには、initial ICWを作成し、適切なビットを設定するだけです。つまり、…

ビット0 - ICW 4を送信するために1に設定します。

x86アーキテクチャには2つのPICがあるので、プライマリPICをスレーブとカスケード接続する必要があります。0のままにしてください。

ビット2 - CALLアドレス間隔。x86では無視され、そのままなので、0にしておきます。

ビット3 - エッジトリガ/レベルトリガンドモードビット。デフォルトでは、エッジトリガになっていますので、0のままにしておきます。

ビット4 - 初期化ビット。1に設定

ビット5...7 - x86では未使用、0に設定されています。

上の図を見ると、最終的なビットパターンは00010001、つまり0x11になります。そこで、PICを初期化するために、ポート0x20にマッピングされているプライマリPICコントローラレジスタに0x11を送ります。

カスケード接続を有効にしているので、ICW 3をコントローラにも送信する必要があります。また、ビット0を設定しているので、ICW4も送信する必要があります。これらについては後ほど説明します。とりあえず、ICW 2を見てみましょう。

ICW 2

このコントロールワードは、PICが使用するIVTのベースアドレスをマッピングするために使用されます。これは重要です。

Initialization Control Word (ICW) 2		
Bit Number	Value	Description
0-2	A8/A9/A10	Address bits A8-A10 for IVT when in MCS-80/85 mode.
3-7	A11(T3)/A12(T4)/A13(T5)/A14(T6)/A15(T7)	Address bits A11-A15 for IVT when in MCS-80/85 mode. In 80x86 mode, specifies the interrupt vector address. May be set to 0 in x86 mode.

初期化の際には、ICW2をPICに送り、使用するIRQのベースアドレスを伝える必要があります。ICW1をPICに送った場合（初期化ビットが設定されている場合）、次にICW2を送る必要があります。そうしないと、定義されていない結果になることがあります。誤った割り込みハンドラが実行される可能性が高くなります。

PICのデータ・レジスタに置かれるICW 1とは異なり、ICW 2はプライマリPICのソフトウェア・ポート0x21、セカンダリPICのポート0xA1として、データ・レジスタに送られます。（PICソフトウェア・ポートの完全なリストについては、「8259Aソフトウェア・ポート・マップ」の表を参照してください）。

さて、先ほどICW 1を両方のPICに送ったと仮定して（上のセクションを参照）、ICW 2を両方のPICに送ってみましょう。これでベースIRQアドレスが両PICにマッピングされます。

これは非常にシンプルですが、PICをどこにマッピングするかに注意しなければなりません。最初の31個の割り込み（0x0-0x1F）は予約されていることを覚えておいてください（上記のx8割り込みベクターテーブル（IVT）表を参照）。そのため、これらのIRQ番号を使用しないようにしなければなりません。

最初の8つのIRQはプライマリPICが処理するので、プライマリPICをベースアドレスの0x20（32進数）に、セカンダリPICを0x28（40進数）にマッピングします。各PICには8つのIRQがあることを覚えておいてください。

簡単でしょう？次の作品に期待しましょう

ICW 3

これは重要なコマンドワードです。PICが相互に通信する際、どのIRQラインを使用するかを知らせるために使用されます。

ICW 3 プライマリPIC用コマンドワード

Initialization Control Word (ICW) 3 - Primary PIC		
Bit Number	Value	Description
0-7	S0-S7	Specifies what Interrupt Request (IRQ) is connected to slave PIC

ICW 3 セカンダリPIC用コマンドワード

Initialization Control Word (ICW) 3 - Secondary PIC		
Bit Number	Value	Description
0-2	ID0	IRQ number the master PIC uses to connect to (In binary notation)
3-7	0	Reserved, must be 0

ICW1の中でカスケードを有効にする時は必ずICW3を送らなければなりません。これにより、どのIRQを使って通信するかを設定することができます。8259Aマイコンは、他のPICデバイスに接続するためにIR0-IR7ピンに依存していることを覚えておいてください。これにより、CAS0-CAS2ピンを使用して相互に通信します。

各PICに互いの情報と、どのように接続されているかを知らせる必要があります。このためには、マスターと関連するPICの両方で使用するIRQラインを含むICW 3を両方のPICに送信します。

覚えておいてください。80x86アーキテクチャでは、IRQライン2を使用してマスターPICとスレーブPICを接続します。

これを知った上で、両方のPICのデータレジスタに書き込む必要があることを忘れずに、上図のフォーマットに従う必要があります。なお、プライマリPICのICW 3では、各ビットが割り込み要求を表しています。つまり...
IRQ Lines for ICW 2 (Primary PIC)

Bit Number	IRQ Line
0	IR0
1	IR1
2	IR2
3	IR3
4	IR4
5	IR5
6	IR6
7	IR7

IRQ2はICW3のビット2ですので、IRQ2を設定するためにはビット2 (0100バイナリ、つまり0x4) を設定する必要があります。ICW 3をプライマリPICに送信する例を示します。

```
; Send ICW 3 to primary PIC
mov al, 0x4           ; 0x4 = 0100 Second bit (IR Line 2)
out 0x21, al          ; write to data register of primary PIC
```

これをセカンダリPICに送るには、2進法で送ることを覚えておかなければなりません。上の表を参照してください。IRQラインを表すには、ビット0~2だけが使われていることに注意してください。バイナリー表記にすることで、8つのIRQラインから選択することができます。

IRQ Lines for ICW 2 (Secondary PIC)	
Binary IRQ Line	
000	IR0
001	IR1
010	IR2
011	IR3
100	IR4
101	IR5
110	IR6
111	IR7

十分にシンプルです。上の表では、2進法<->12進法の会話に従っているだけであることに注意してください。

IRQライン2で接続されているため、ビット1を使用する必要があります（上図参照）。

ここでは、プライマリとセカンダリの両方のPICコントローラにICW 2を送信する完全な例を示します。

```
mov al, 0x4           ; 0x04 => 0100, second bit (IR line 2)
out 0x21, al          ; write to data register of primary PIC

; Send ICW 3 to secondary PIC
mov al, 0x2           ; 010=> IR line 2
out 0xA1, al          ; write to data register of secondary PIC
```

それが全てです。)

さて、ここで両方のPICがIRライン2を使って通信するように接続されました。また、両方のPICが使用するベース割り込み番号を設定しました。

これは素晴らしいことですが、まだ終わりではありません。ICW1をビルトアップする際、ビット0がセットされていれば、PICはICW4を送ることを期待していることを覚えておいてください。そのため、最終的なICWであるICW4をPICに送る必要があります。

ICW 4

イェーイ！これは、最終的な初期化コントロールワードです。これは、すべての動作を制御します。

Initialization Control Word (ICW) 4

Bit Number	Value	Description
0	uPM	If set (1), it is in 80x86 mode. Cleared if MCS-80/86 mode
1	AEOI	If set, on the last interrupt acknowledge pulse, controller automatically performs End of Interrupt (EOI) operation
2	M/S	Only use if BUF is set. If set (1), selects buffer master. Cleared if buffer slave.
3	BUF	If set, controller operates in buffered mode
4	SFNM	Special Fully Nested Mode. Used in systems with a large amount of cascaded controllers.
5-7	0	Reserved, must be 0

これはかなり強力な関数です。5~7ビット目は常に0なので、他のビットやピースに注目してみましょう。

PICは、80x86が登場する以前から、汎用的なマイクロコントローラとして設計されてきました。そのため、システムに合わせてさまざまな動作モードが用意されていますが、その中のひとつに「特殊フルネストモード」というものがあります。

x86ファミリーはこのモードをサポートしていないので、ビット4を0に設定しておくと安心です。

ビット3はバッファードモードに使用します。動作モードについては後で説明するので、とりあえず0にしておきます。ビット2は、ビット3が設定されているときにのみ使用されるので、

これを0に設定します。これにより、ビット1もほとんど使用されません。

そのため、ビット0を設定するだけで、PICの80x86モードが有効になります。簡単ですね。つまり、ICW 4を送るために必要なのは、次のようにになります。

```
mov al, 1           ; bit 0 enables 80x86 mode
; send ICW 4 to both primary and secondary PICs
out 0x21, al
out 0xA1, al
```

これは、このチュートリアルの中で最も簡単なコードスニペットでしょう。使えるうちに使ってみてください。:)

PICを初期化する - まとめてみる

信じられないかもしれません、これについてはすでに説明しました。PICの初期化では、正しいICWをPICに送るだけでいいのです。

ここでは、前のセクションのすべてをまとめてPICを初期化し、すべてがどのように組み合わされているかを理解しましょう。

```
;*****
8259A PICで割り込みテーブルの32-47を使用するようにマッピングします。
;*****

#define ICW_1 0x11          ; 00010001 binary. Enables initialization mode and we are sending ICW 4
#define PIC_1_CTRL 0x20      ; Primary PIC control register
#define PIC_2_CTRL 0xA0      ; Secondary PIC control register
#define PIC_1_DATA 0x21      ; Primary PIC data register
#define PIC_2_DATA 0xA1      ; Secondary PIC data register
#define IRQ_0   0x20          ; IRQs 0-7 mapped to use interrupts 0x20-0x27
#define IRQ_8   0x28          ; IRQs 8-15 mapped to use interrupts 0x28-0x36

MapPIC:

Send ICW 1 - Begin initialization -----
; Setup to initialize the primary PIC. Send ICW 1
mov al, ICW_1
out PIC_1_CTRL, al

Send ICW 2 - Map IRQ base interrupt number -----
; Remember that we have 2 PICs. Because we are cascading with this second PIC, send ICW 1 to second PIC command register
out PIC_2_CTRL, al

ICW 2をプライマリPICに送信
mov al, IRQ_0
out PIC_1_DATA, al
; send ICW 2 to secondary controller
mov al, IRQ_8
out PIC_2_DATA, al

; Send ICW 3 - 両方のPICを接続するためにIRラインを設定する -----
ICW 3をプライマリPICに送る
mov al, 0x4
out PIC_1_DATA, al          ; 0x04 => 0100, second bit (IR line 2)
                           ; write to data register of primary PIC

ICW 3を2次PICに送る
mov al, 0x2
out PIC_2_DATA, al          ; 010=> IR line 2
                           ; write to data register of secondary PIC

ICWの送信 4 - x86モードの設定
mov al, 1                   ; bit 0 enables 80x86 mode
; send ICW 4 to both primary and secondary PICs
out PIC_1_DATA, al
out PIC_2_DATA, al
; All done. Null out the data registers
mov al, 0
out PIC_1_DATA, al
out PIC_2_DATA, al
```

それほど難しくなかったでしょう？このコードですべてをカバーしました。

これでPICの初期化が完了しました。ハードウェア割り込みが発生すると、あらかじめIVT（Interrup Vector Table）のどこかに定義しておいた32～47番の割り込みが呼び出されます。

これにより、ハードウェア割り込みを追跡することができます。かっこいいでしょう？

オペレーション・コマンド・ワード (OCW)

うれしいですね。さて、醜い初期化作業が終わったところで、いよいよPICの標準的な制御と操作に集中することができます。そのためには、OCW (Operation Control Words) を使って、さまざまなレジスタの書き込みや読み出しが行います。

OCW 1

OCW 1は、IMR (Interrupt Mask Register) 内の値を表す。現在のOCW 1を知るには、IMRから読めばよい。

IMRは、ステータスレジスタと同じポートにマッピングされていることを覚えておいてください。ステータスレジスタはリードオンリーなので、PICではリードかライトかでアクセスするレジスタを決定しています。

前述のPICレジスタの説明の際にIMRレジスタを見ました。

OCW 2

これは、PICを制御するための主要なコントロールワードです。それでは見てみましょう。

Operation Command Word (OCW) 2		
Bit Number	Value	Description
0-2	L0/L1/L2	Interrupt level upon which the controller must react
3-4	0	Reserved, must be 0
5	EOI	End of Interrupt (EOI) request
6	SL	Selection
7	R	Rotation option

じゃあ、いいか！？ビット0～2は、現在の割り込みの割り込みレベルを表します。3-4ビットは予約済みです。5-7ビットは、興味深いビットです。これらのビットの各組み合わせを見てみましょう。

OCW2 Commands			
R Bit	SL Bit	EOI Bit	Description
0	0	0	Rotate in Automatic EOI mode (CLEAR)
0	0	1	Non specific EOI command
0	1	0	No operation
0	1	1	Specific EOI command
1	0	0	Rotate in Automatic EOI mode (SET)
1	0	1	Rotate on non specific EOI
1	1	0	Set priority command
1	1	1	Rotate on specific EOI

さて…この表、今までは分かりにくいと思いませんか？上記のコマンドの多くはかなり高度なものです。では、どんなことができるのか見てみましょう。

EOI (End of Interrupt) の送信

ご存知のように、ハードウェア割り込みが発生すると、プライマリコントローラにEOI信号が送られるまで、他のすべての割り込みは割り込みマスクレジスタ内でマスクされます。

つまり、割り込みルーチン (IR) の最後に、すべてのハードウェア割り込みが有効になるようにEOIを送信する必要があります。

上の表を見ると、コントローラにEOIの信号を送るために、特定ではないEOIコマンドを送ることができます。EOIビットはOCW2のビット5なので、ビット5 (100000バイナリ = 0x20) を設定するだけでよい。

```
mov    al, 0x20      ; set bit 4 of OCW 2
out   0x20, al      ; write to primary PIC command register
```

Conclusion

PICはプログラムが複雑なマイクロコントローラです。このチュートリアルでは多くのことを説明してきましたが、さらに難しくなってきています。

うまく説明できたでしょうか？OS開発シリーズのプライマリ・チュートリアル・シリーズでは、このチュートリアル内のすべての内容を、あるべき場所に配置します。これは、割り込みの設定、割り込み処理、ハードウェア割り込みの間の接着剤となるでしょう；)

このチュートリアルは、より多くの内容を提供し、8259Aマイクロコントローラに関するすべての詳細を説明できるように拡張する予定です。

このチュートリアルは、本編が使用するサイドチュートリアルです。そのため、このチュートリアルのためのデモはありません。しかし、すべてをまとめるために、デモを作ることになるかもしれません。ただし、そのためにはIDTと割り込み処理を詳しく説明する必要がありますが、これはPICには関係ありません。直接的には関係ありませんが。

このチュートリアルが、PICをプログラミングする際の疑問や、ボンネットの中で実際に何が起こっているのかを理解する助けになれば幸いです。それではまた次回、お会いしましょう。

~マイク

*BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中です。質問やコメントはありますか？お気軽にお問い合わせください。
あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。*

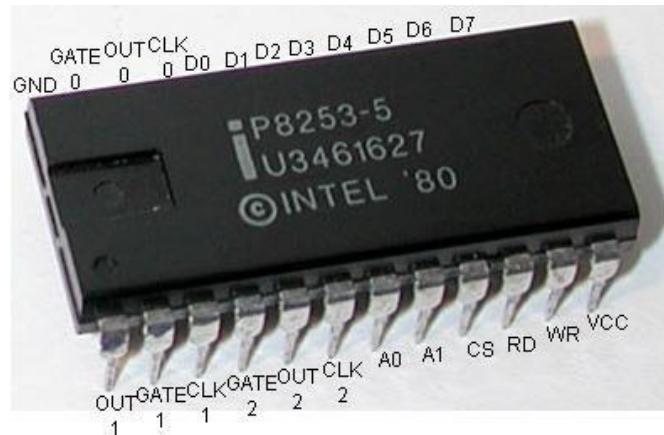


オペレーティングシステム開発シリーズ

オペレーティング・システムの開発 - 8253プログラマブル・インターバル・タイマー

by Mike, 2008

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。



8253 PIT マイクロコントローラ (全ピン表示付き)

ご注意：このチュートリアルでは、ハードウェア割り込み処理や8259 Programmable Interrupt Controller (PIC)に関する知識が必要になる場合があります。これらの情報については、このチュートリアルをご覧ください。

はじめに

歓迎します。:)

このチュートリアルでは、システム・タイミングとIntel 8253 Programmable Interval Timer (PIT)のプログラミングについて知りたいことをすべて説明します。

8253 PITには長い歴史があり、ほぼすべてのx86 PCで重要な役割を果たしています。これは「システムクロック」と呼ばれ、PCの中で非常に重要な機能を担っています。この... 「このチップは、もはや独立したチップとして（正確にはDIP (Dual Inline Package) として）流通しておらず、むしろマザーボードのサウスブリッジ・チップセットに組み込まれています。

しかし、8253のすべてが残っています。そのため、入出力設備、ハードウェア、そして8253のプログラミング方法も同じです。このDIPと旧式のDIPの間には、速度以外の違いはないので、ここではシンプルにするために旧式の8253のDIPを見ていきます。

このチュートリアルの冒頭の写真は、これから見ていくもの、プログラミングするものを表示しています。
楽しくやりましょう。)

プログラマブル・インターバル・タイマー

プログラマブル・インターバル・タイマー(PIT)は、プログラムされたカウントに達すると、割り込みを発生させるカウンタです。8253および8254マイクロコントローラは、i86アーキテクチャに対応したPITで、i86対応システムのタイマとして使用されます。

これらのPITには、異なる目的で使用される3つのタイマーが含まれています。最初のタイマーは通常、システムクロックとして使用されています。タイマー2はRAMのリフレッシュに使われ、タイマー3はPCのスピーカーに接続されています。すべての接続を見るのはもう少し後になりますので、今はあまり詳しく説明しません。

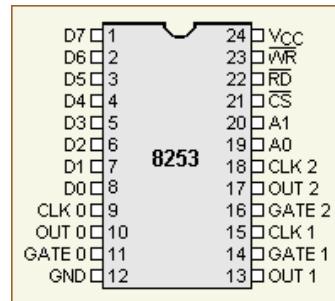
ここでは、そんな有名なPITのひとつを紹介しましょう...。8253マイクロコントローラ。

8253ハードウェア

ソフト面を見る前に、実際に何をプログラムしているのかを知っておくと便利です。そこで今回は、まず8253のハードウェアを取り上げ、その動作とPCの他の部分との接続について学びます。また、ソフトウェア側で必要となる内部レジスタ、ピン配置、コマンドワードなどについても見ていきます。

8253ハードウェア。説明

8253 PITはシンプルなインターフェイスで、プログラミングもそれほど難しくありません。



そうです。私たちがプログラムするチップがあるのです。

これがピンの完全なレイアウトです。このチュートリアルでは、これらのピンを参照しますので、ピンが何であるかを知っておくことが重要です。

D0...D7 : 8ビットのデータライン。データバスに接続されているので、コマンドの読み出しや送信が可能です。

CLK 0, CLK 1, CLK 2: クロック入力端子。3つの独立したカウンター用の3端子です。OUT 0,

OUT 1, OUT 2 : 出力データラインです。GATE 0, GATE 1, GATE 2: ゲートデータライン。3つの独立したカウンター用の3つの端子があります。

Vcc : 入力電圧

WR : ライトイネーブル。このラインがアクティブになると、8253にデータを書き込ん

でいることを知らせます。 RD: リードイネーブル。この信号がアクティブになると、8253は

データを読み出すことができる CS: チップセレクト信号

A0、A1 : アドレスライン。どのレジスタにアクセスしているかを判断するために使用する。

悪くないですね。ここには見ておくべき重要なピンがいくつかある。

D0～D7ピンは、システムのデータバスに接続します。これらのピンは、コントローラにデータを送信または読み取る際に、データを運びます。

VccとGNDで回路が完成します（電圧入力、グランド出力）。

WRピンは、コントローラに書き込み中であることを伝えます（データピンへの入力を期待します）。このピンの信号が "Low" のとき、現在データを送信しています。RD端子も同様ですが、データを読み込んでいることをコントローラに伝えます。CSピンは、コントローラがRDピンとWRピンに対して何をすべきかを決定する特別なピンです。CSピンが "Low" であれば、コントローラはRDとWRピンに反応します。CSが "Low" でない場合、それらは無視されます。WRとRDはシステムコントロールバスに接続します。CSピンは、ポートi/o操作のためにシステム・アドレス・バスに接続します。

A0、A1ピンはシステムアドレスバスに接続されており、どのレジスタにアクセスしているかを判断するために使用されます。A0、A1ピンはシステムアドレスバスに接続されており、どのレジスタにアクセスしているかを判断するために使用されます。

CLK端子、OUT端子、GATE端子の3つのグループがあることに注目してください。これにはちゃんと理由があります。8253/8254マイコンには、3つの独立したタイマーが搭載されています。

この仕組みをもう少し詳しく見てみましょう。

8253のハードウェア。カウンター

8253は、3つのカウンタで構成されています。カウンタ0、カウンタ1、カウンタ2です。各カウンタには2つの入力端子があります。CLK(クロック入力)

とGATEの2つの端子と、出力用の1つの端子--OUT。

カウンタは3つあるので、システム内でそれぞれ異なる目的で使用されます。各カウンタは16ビットのダウンカウンターです。

一般的なコンピュータでは、第1タイマのOUTピンをPIC (Programmable Interrupt Controller) に接続して、クロックの刻みごとに割り込みを発生させます。これは通常、システムタイマーとして使用されます。2番目のカウンタは、DRAMメモリをリフレッシュするためにメモリコントローラにタイミング信号を生成するために使用された。3番目のカウンタは、PCのスピーカーに音を出すために使用されます。

お察しの通り、PITはカウンタが0になるとOUTピンを使ってこれらのデバイスに信号を送ります。PITのカウンタが0になると、単純に折り返して再スタートします。

CLKは、タイマーのクロック入力です。現在の動作モードに応じて、GATE端子と併用することができます。以下の表は、GATEに流れる電流がLow, Rising, Highのいずれかである場合の動作を示しています。

GATE Input pin operations			
Mode	Low or going low	Rising	High
0	Disables Counting	-	Enables Counting
1	-	Initiates Counting and resets OUT after next CLK	-
2	Disables Counting, Sets OUT to high	Reloads counter and initiates counting	Enables Counting
3	Disables Counting, Sets OUT to high	Initiates Counting	Enables Counting
4	Disables Counting	-	Enables Counting
5	-	Initiates Counting	-

8253のカウンターはチャンネルと呼ばれています。8253/8254 PITには3つのチャンネルがありますが、それぞれのチャンネルについてもう少し詳しく見てみましょう…

チャンネル0

チャンネル0は、8259 PICに接続され、割り込み要求 (IRQ) を生成します。PITs OUTピンはPICのIR0ピンに接続します。通常、BIOSはこのチャンネルを65536カウントで構成し、出力周波数は18.2065Hzになります。これにより、54.9254msごとにIRQ 0が発生します。

ほとんどのx86マシンで使用されている主要なタイマーです。クロックレート（カウンター0のCLKピンで通知される）は1193181.6666...Hzです。Hzで、これはNTSCサブキャリア周波数の3分の1にあたります。これは、旧来のCGA PCとの互換性を保つために必要なものである。

多くのシステムでは、チャンネル0がシステムクロックとして動作するようにプログラムされています。これは、チャンネル0のOUTピンがPICのIR0ラインに間接的に接続されているために可能です。設定するモードに応じて、タイマーを適切な周波数に設定し、PICのIR0ラインを一定の割合で効果的にすることができます。その後、リセットされて、また最初からやり直しです。PICはハードウェア割り込みを扱うため、まずPICを再プログラムする必要があります。

また、最も番号の小さい割り込み (IRQ0) に接続されているため、他のハードウェア割り込みよりも優先順位が最も高くなっています。

最も低い周波数は、昔のDOSシステムを搭載したコンピュータに使われていたもので、約18.2Hzです。最高周波数は1メガヘルツ強。リアルモードのOSでは、BIOSは通常、IRQ0の発火回数を 0000:046C にインクリメントし、実行中のプログラムがこれを読み取ることができます。

チャンネル1

多くのビデオカードやBIOSは、独自の用途のために第2チャンネルを再プログラムすることができます。このチャンネルは元々、メモリコントローラがDRAMメモリをリフレッシュするためのタイミングパルス信号を生成するために使用されていました。現代では、メモリコントローラがリフレッシュを行うため、この信号は必要ありません。このため、どのような機器でこのカウンターが使用されるかは保証されていません。

チャンネル2

このチャンネルは、PCスピーカーに接続して音を発生させます。PCスピーカーは通常、2つのレベルの出力を持つ矩形波を生成するようになっています。しかし、真の意味で定義された2つの矩形波レベルの間を行き来することが可能です。これをPWM(Pulse-Width Modulation)といいます。

このチャンネルの設定は、モード3にプログラムし、トーンの周波数レートを設定します。

PCスピーカーを直接プログラムすることもできます。チュートリアル7を振り返ると、PCスピーカーがポート0x61にマッピングされていることがわかります。このポートは、スピーカーがどのように動作するかを定義します。

ビット0：(1)に設定すると、スピーカーの状態はビット1に従う

ビット1：設定されている場合(1)、スピーカーはPITを使用し、設定されていない場合(0)、スピーカーはPITとの接続を無効にする

ビット0がセットされていれば、残りのバイトには音の周波数を表すビットのパターンが含まれています。スピーカーから最大8ビットの音を発生させることができます。これはちょっとかっこいいですね。

また、PITを使用しないようにデバイスを無効化することにも注目してください。起動時に、BIOSはスピーカーがPITチャンネル2を使用し、モード3で動作するように設定します。タイミングの問題が発生する可能性があるため、スピーカーはPITを使用するようにしておくことをお勧めします。念のため、例を挙げておきます。

```
; disables the speaker, and stop using channel 2
mov    dx, 0x61
out   dx, 0

; generates tone from speaker
out   dx, 11111101b
```

結論

一度設定されたカウンターは、他のコントロールワードで変更されるまで、その状態が維持されます。

このカウンターを使って、何か面白いことができそうですね。チャンネル1はもう使われていないので、安全だと思って自分で使うことはできません。そのため、チャンネル0と2を使うことをお勧めします。

チャンネル0を使って、割り込みハンドラを起動することができます。割り込みハンドラは、カーネルが使用するカウンタをインクリメントすることができます。この特別な小さなカウンタ変数は、システムにおいて非常に重要な役割を果たしています。システムタイマーです。これらのこととはすぐにわかるでしょう。)

さて、ここまでではピンの構成と、異なるデバイスで使用される3つのタイマーについて見てきました。次は何ですか？

これらのタイマーをプログラミングする際には、初期化を行う必要があります。各チャンネルは6つの異なるモードをサポートしていることを覚え

ておいてください。これらのモードの中には非常に便利なものもありますが、そうでないものもあります。他のモードはそうではありません。これらのモードを理解するために、それぞれのモードを見てみましょう。少し詳しくなりますが、すでにご存じの方、あるいは期待されている方もいらっしゃると思いますので、ご了承ください。)

8253のチャンネルモード

各カウンターは、6つのモードのうち1つにプログラムできることを覚えておいてください。これを行うには、コントローラに初期化制御ワード (ICW) を送信します。このコマンドワードのフォーマットについては後ほど説明します。ここでは、各モードについて説明します。

モード0：端子カウントでのインタラプト

このモードでは、カウンタは最初のCOUNT値にプログラムされ、その後、入力クロック周波数 (CLK信号) に合わせてカウントダウンしていきます。COUNTが0のとき、コントロールワードが書き込まれた後、カウンタは、接続されているデバイスに信号を送るためにOUTピンをイネーブルにします（ラインをハイにします）。カウントは、COUNTがプログラムされた1クロックサイクル後に開始されます。OUTラインは、カウンタが新しい値または同じ値で再ロードされるか、別のコントロールワードがコントローラに書き込まれるまで、ハイレベルのままでです。このモードは基本的に、0までカウントダウンするタイマーを設定することができます。その後、新しいカウント番号を再ロードするか、カウンターを再初期化するために新しいコントロールワードを再ロードする必要があります。

モード1：ハードウェアトリガーによるワンショット

このモードでは、カウンタは一定のクロックパルス数ごとに出力パルスを与えるようにプログラムされています。コントロールワードが書き込まれると、すぐにOUTラインがハイレベルになります。COUNTが書き込まれた後、カウンタはGATE入力の立ち上がりエッジまで待機します。パルス出力中にトリガが発生した場合、8253は再び再トリガされます。GATEの立ち上がりエッジが検出されてから1クロック後にOUTはLOWになります、COUNTが0になるまでLOWを維持します。その後、OUTは次のトリガがかかるまでHIGHになり、GATE入力の立ち上がりエッジが検出されるまで再び待機します。

モード2：レートジェネレーター

このモードでは、カウンターを "divide by n" カウンターに設定します。このカウンターは、一般的にリアルタイムシステムクロックの生成に使用されます。カウンタは、初期のCOUNT値にプログラムされます。カウントは次のクロックサイクルで開始されます。OUTは、COUNTが0になるまでHighのままでです。

は1に達します。その後、1クロックパルスの間、OUTはLOWになります。その後、OUTは再びHighになります、COUNTは初期値にリセットされます。このプロセスは、新しい制御ワードがコントローラに送られるまで繰り返されます。ハイパルス間の時間は、「COUNT」の現在値に依存し、次の式で算出されます。

```
COUNT = input (Hz) / Frequency of output
```

COUNTが0になることはなく、nから1までの範囲でしかありません (nはCOUNTの初期値)。

さて、ちょっと立ち止まってみましょう。カウンタ0がPICに接続されていることを覚えていますか？カウンタ0のOUTラインは、間接的にPICのIRQ0ラインに接続されています。これを知っていると、IRQ0ラインがLowの時、PICは我々が定義したIRQ 0ハンドラを呼び出します。

カウンタをモード2に設定すれば、一定の割合で割り込みが発生するようにタイマーを設定することができます。あとは、上の式をもとにCOUNTの値を決めるだけです。これは、OSのシステムタイマーの設定によく使われます。結局のところ、IRQ 0は、定義した周波数レートで、クロックティックごとに呼び出されることになります。

確かに、モード2は重要なモードですね。

モード3：スクウェア・ウェーブ・ジェネレータ

このモードは、モード2とよく似ています。ただし、OUTは期間の半分はHigh、残りの半分はLowになります。COUNTが奇数の場合、OUTは(n+1)/2カウントの間ハイになります。COUNTが偶数の場合は、(n-1)/2カウントの間、OUTはLowになります。

それ以外はモード2と同じです。COUNTの初期値を設定するには、モード2の計算式を使う必要があります。スピーカーがPITを使用するように設定されている場合は、通常使用するチャンネルをこのモードを使用するように設定する必要があります。

モード4：ソフトウェア・トリガ・ストローブ

カウンタは、初期のCOUNT値にプログラムされています。カウントは次のクロックサイクルで開始されます。OUTは、COUNTの値が大きくなるまでHighのままでです。

が0になると、カウンタは1クロックの間、OUTをLowにします。その後、再びOUTをHighにリセットします。

モード5：ハードウェア・トリガ・ストローブ

カウンタは、初期のCOUNT値にプログラムされています。OUTは、コントローラがGATE入力の立ち上がりエッジを検出するまでHighのままでです。これが起こると、カウントが開始されます。COUNTが0になると、1クロックサイクルの間、OUTはLOWになります。その後、OUTは再びハイレベルになります。このサイクルは、コントローラがGATEの次の立ち上がりエッジを検出するまで繰り返されます。

8253レジスター

8253には、アクセスできるレジスタがいくつかあります。これらのレジスタのほとんどは互いによく似ているので、わかりやすくするために同じ表にまとめておきます。この表は、8253の対応するラインがアクティブなときの各レジスタとその機能を示しています。RDラインとWRラインが読み出しと書き込みの動作を決定することに注目してください。また、A0とA1のラインが、どのレジスタにアクセスしているかを決定していることにも注目してください。

チートリアル7のポートテーブルを見ると、システムタイマがBIOSによってポート0x40-0x4Fを使用するようにマッピングされていることがわかります。各ポートのアドレスは1バイトの大きさです。

8253 PIT Internal Registers						
Register Name	Port Address	RD line	WR line	A0 line	A1 line	Function

Counter 0	0x40	1	0	0	0	Load Counter 0
		0	1	0	0	Reads Counter 0
Counter 1	0x41	1	0	0	1	Load Counter 1
		0	1	0	1	Reads Counter 1
Counter 2	0x42	1	0	1	0	Load Counter 2
		0	1	1	0	Reads Counter 2
Control Word	0x43	1	0	1	1	Write Control Word
NA		0	1	1	1	No Operation

それ以外の0x44～0x4fのポートアドレスは未定義です。

システムは、実行している操作に応じて、適切なラインをアクティブにします。カウンタレジスタを設定する際には、まずどのようにロードするのかをコントローラに知らせる必要があります。これは、最初にコントロールワードを設定することで行われます。それでは、これらのレジスターを詳しく見てみましょう…

カウンタレジスター

各カウンタレジスタには、PITがカウントダウンに使用するCOUNT値が格納されています。これらはすべて16ビットのレジスタです。これらのレジスタに書き込みや読み出しを行う場合は、まずPITに制御ワードを送る必要があります。なぜそれを直接できないのかと思われるかもしれません。これには理由があって、データの大きさに関係しています。PITには8本のデータライン（ピンD0～D7）しかありません。しかし、カウンタレジスターは8ビットではなく、すべて16ビットです。

このため、PITはカウンタレジスタに書き込まれたデータをどのようにして知るのでしょうか？カウンタレジスタの16ビットの中のどのバイトを設定しているのか、どうやって知っているのでしょうか？そうではありません。コマンド・ワードを送信することで、PITにデータが入ってくること、そしてそのデータをどうするかを知らせることができます。これについては次に説明します。

コントロール・ワード・レジスター

これは私たちにとって重要なことです。

このレジスタは、コントローラの動作モードを決定し、設定するために使用される重要なレジスタです。このレジスタにアクセスするには、RD, A0, A1ラインをイネーブルにする必要があります。このレジスタは書き込みのみ可能で、読み出しができません。

コントロール・ワード・レジスターは、シンプルなフォーマットを使用しています。最初は表にしようかと思っていたのですが、リスト形式の方が簡単かもしれないのに、ここで紹介します。

ビット0: (BCP) バイナリカウンタ

0: バイナリ

1: BCD (バイナリコード化された10進法)

ビット1～3: (M0, M1, M2) 動作モード。それぞれの説明は上記のセクションを参照してください。

000: モード0: インタラプトまたはターミナルカウント

001: モード1: プログラマブルワンショット

010: モード2: レートジェネレータ

011: モード3: 矩形波ジェネレータ

100: モード4: ソフトウェア・トリガ・ストローブ

101: モード5: ハードウェア・トリガ・ストローブ

110: 未定義、使用しないでください

111: 未定義、使用しない

ビット4-5: (RL0, RL1) リード／ロードモード。カウンタレジスタにデータを読み込んだり、送信したりします

00: カウンタ値は、I/O書き込み動作時に内部制御レジスタにラッチされます。

01: 最下位バイト (LSB) のみ読み出しましたは読み込み 10:

最上位バイト (MSB) のみ読み出しましたは読み込み 11:

LSBの次にMSBを読み出しましたは読み込み

ビット6-7: (SC0-SC1) セレクトカウンター。それぞれの説明は上記のセクションを参照してください。

00: カウンタ0

01: カウンタ1

10: カウンタ2

11: 不正な値

さて、それでは！ちょっとしたことをやってみましょうか。あとは、コントロール・ワード・レジスターに書き込んで、コントロール・ワードを構築し、コントローラを初期化するだけですね？もちろんだよ。そうだな…。

基本的には、特定の目的のためにカウンターを初期化したい。そのため、カウンター、カウンターのカウントモード、動作モードを設定するためのコントロールワードを構築する必要があります。その後、カウンター自体を初期化します。一度初期化すると、カウンターは（そのモードに応じて）それ自体を継続することができるので、この作業は一度だけ行う必要があることを覚えておいてください。

一例を挙げて、すべてをまとめてみましょうか。

すでにPICを初期化し、割り込み0のハンドラがあるとします。100Hz (10ミリ秒に1回) ごとにIRQ 0を発生させるタイマーをセットアップしたいとします。PITのチャンネル0がPICのIR0ラインに接続されていることがわかっているので、チャンネル0をプログラムしてこれを実行します。

```
; COUNT = input hz / frequency  
mov     dx, 1193180 / 100      ; 100hz, or 10 milliseconds
```

```

mov      ax, dx
out     0x40, al      ;LSB
xchg    ah, al
out     0x40, al      ;MSB

```

最初にコントロールワードを設定し、次にカウンタ0レジスタに書き込んでいることに注目してください。

そうなんですか！？うん。以上により、カウンタ0が10ミリ秒ごとにIRQ0を起動するようにプログラムされます。

結論

8253と8254のPITは、非常に使い勝手の良い小さなチップです。様々なデバイスに使用でき、様々な目的に使用できます。

私たちが必要とするのは、PCのスピーカーからの信号出力と、最新のシステムソフトウェアでは非常に重要な機能であるシステムタイマーの両方を実現することです。さらに、スピーカーを直接操作する方法や、PITとの接続を無効にする方法も検討しましたが、これにはメリットとデメリットがあります。システムソフトウェアの設計者は、システムを開発する際に、さまざまな無限の可能性の中から長所と短所を判断し、それを両立させることができます。

ここまで、標準的なx86ベースのPCマザーボードに搭載されているPITの詳細、ピン配置、およびその接続について見てきました。最近のパソコンでは、PIT自体がマザーボードのサウスブリッジと一体化しています。

私は、チップ自体にさらに追加して、一般的なコンピュータシステムの中でチップがどのように接続されているかを正確に説明しようと考えています。いっそのこと、分解してしまってもいいかもしれません。

チュートリアル 16: Kernel: タイミングと例外処理では、8259A PICとのチュートリアルのすべてをまとめます。両方のデバイスを実装し、インターフェースを作成します。もしかしたら、もうちょっと…。

Welcome back to kernel Land!

次の機会まで。

~マイク

BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中です。質問やコメントはありますか？お気軽にお問い合わせください。
あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。



オペレーティングシステム開発シリーズ

オペレーティングシステム開発 - グラフィックス1

by Mike, 2010

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

はじめに

Welcome!

待って、何？もうグラフィック？そうなんです、OSの超クールなGUIの開発を始めますよ。:) そうではありませんが、その方向に向けてのスタートです。

本章は、グラフィックスプログラミングを扱うミニシリーズの最初の章です。この章では、VESA VBE、Video BIOS、VGA用の直接ハードウェアプログラミング、そしておそらくSVGAのコンセプトを取り上げます。また、2Dベクトルレンダリングや画像など、グラフィックの概念やレンダリングについても取り上げる予定です。もしかしたら、もう少し後には3Dも扱うかもしれません。

興奮していますか？OS開発シリーズのスピンオフであるこのミニシリーズには、たくさんのクールな素材が登場します。しかし、素晴らしいコンピューターグラフィックスの世界に飛び込む前に、基本的なルールを決めておかなければなりません。一口にCGといっても、その扱い方や方向性はさまざまです。コンピュータ・グラフィックスは複雑なテーマです。1つの章ではカバーできません。いや、できるんですよ。ただ、1つの.....とてもとても長い章になってしまいますが。

そのため、これを段階的に行うことにしました。最初の章では、リアルモードやv86モードでのグラフィックの扱いについて説明します。システムBIOSの割り込みを使い、グラフィックスの基本的な概念を説明します。第2章では、「Video BIOS Extensions (VBE)」と「Super VGA」を取り上げます。第3章では、グラフィックス・パイプラインの直接ハードウェア・プログラミングを取り上げた、より小さなミニシリーズの第1章となります。第3章では、グラフィックス・パイプラインのダイレクト・ハードウェア・プログラミングを取り上げます。

この章では、まずリアルモードのVideo BIOSを使って、リアルモードのグラフィックスを操作してみましょう....。

基本コンセプト

アブストラクト

コンピュータ・グラフィックス (CG) は、今さら説明するまでもありません。コンピュータ、アニメーション、ビデオゲーム業界に革命をもたらしました。コンピュータグラフィックスの分野は、コンピュータのディスプレイ上でグラフィック効果を生み出す能力の開発、創造、継続を網羅しています。1Dグラフィックスから、2D、3D、そして4Dグラフィックスのシミュレーションソフトまで。

歴史

コンピュータグラフィックス産業は、1960年代の「ワールウィンド」のような初期のプロジェクトから生まれた。

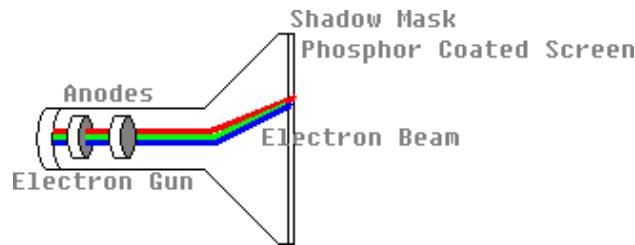
Whirlwindは、ビデオディスプレイ出力を使用した最初のコンピューターであり、CRT (Cathode Ray Tube) 技術の導入にも貢献した。Whirlwindは、最終的にSAGE (Air Force Semi Automatic Ground Environment) コンピュータシステムの開発につながった。ブラウン管は、1897年にフェルディナンド・ブラウンが開発した「ブラウン管」が最も古いものである。

SIGGRAPHグループは、世界各地でSIGGRAPHカンファレンスを開催しています。これらのカンファレンスには、エンジニアリング、グラフィックス、映画、ビデオゲーム業界の企業から何千人ものプロフェッショナルが参加しています。

グラフィックス・ハードウェアの進歩により、より強力なグラフィックス・デザインが可能になった。液晶ディスプレイ(LCD)などの他のディスプレイ技術が登場すると、CRT技術の使用は減少していった。

VDT (Video Display Terminals) は、VDU (Video Display Unit) とも呼ばれる初期のディスプレイ端末。

Cathode ray tube (CRT)



Abstract

CRTは、電子銃と蛍光体ターゲットで構成された真空管である。ブラウン管の前面全体をラスターと呼ばれるパターンで繰り返し走査する。画像は、赤、緑、青の3色の電子ビームの強度をディスプレイ上の任意の位置で変化させることで生成される。電子ビームは、シャドウマスク層を通過した後、蛍光体を塗布したスクリーンに当たる。

問題点

CRTモニターは微量のX線を放出することがあります。また、ディスプレイを常に再スキャンしているため、低リフレッシュレート（60Hz以下）ではちらつきが見られることがあります。また、CRTには有害な蛍光体が含まれている場合があります。このため、米国環境保護庁（EPA）は、CRTを適切なリサイクル施設に買い取らなければならないという規則を作りました。最後に、CRTには真空のガラスが入っているため、外側のガラスが破損すると、CRTが破裂する可能性があります。これは、ガラスが危険な速度で外側に飛び散る原因となります。最近のCRTは、CRTの飛散を防ぐために一定の対策が施されている。

CRTの周波数をソフトウェアで制御することが可能です。周波数が高くなると、CRTを本来の用途よりも高速に動作させることが可能となり、CRTが破裂する可能性が高くなります。このため、CRTコントローラ（CRTC）の取り扱いには十分な注意が必要です。しかし、最近のCRTはこのようなことがないように保護されています。

VGA

Abstract

VGA（Video Graphics Array）は、1987年にIBMが発売したアナログコンピュータディスプレイ規格。「アレイ」と呼ばれるのは、もともとMDA、CGA、EGAが使用していたISA（Industry Standard Architecture）ボードの数十個のロジックチップに代わる1つのチップとして開発されたからである。ISAボード1枚に収まっていたため、マザーボードへの接続が非常に容易でした。

VGAは、ビデオバッファ、ビデオDAC、CRTコントローラ、シーケンサユニット、グラフィックスコントローラ、属性コントローラで構成されています。これらの構成要素については、後の章で詳しく説明します。

ビデオバッファ

ビデオバッファは、ビデオメモリとしてマッピングされたメモリのセグメントです。メモリのどの領域がビデオメモリにマッピングされるかは変更できる。起動時には、BIOSは0xA0000にマッピングします。つまり、ビデオメモリは0xA0000にマッピングされます。（チュートリアル7のリアルモードアドレスマップを覚えていますか？）メモリマッピングについては、この章の少し後に詳しく説明します。

ビデオDAC

Video Digital to Analog Converter (DAC)には、ビデオデータをディスプレイに送るアナログビデオ信号に変換するためのカラーパレットが含まれています。この信号は、赤、緑、青の色の濃さをアナログで表したもので、詳しくは後で説明しますので、まだ理解できなくてもご安心ください。

CRTコントローラー

このコントローラは、水平・垂直同期信号のタイミング、ビデオバッファのアドレッシング、カーソルやアンダーラインのタイミングを生成する。詳しくは、後ほどVGAハードウェアの説明の際に説明します。

シーケンサー

シーケンサは、ビデオメモリの基本的なメモリタイミングと、再生バッファのフェッチを制御するためのキャラクタクロックを生成します。これにより、システムはアクティブなディスプレイのインターバル中にメモリにアクセスすることができます。もう一度言いますが、これについてはまだ詳しくは説明しません。

グラフィックスコントローラ

ビデオメモリとアトリビュートコントローラ、ビデオメモリとCPUの間のインターフェースである。表示がアクティブな時間帯には、ビデオバッファ（ビデオメモリ）からメモリデータが送られ、アトリビュートコントローラに送られる。グラフィックスモードでは、このデータをパラレルからシリアルのビットプレーンデータに変換してから送信する。テキストモードでは、パラレルデータだけが送信されます。

まだ理解できなくても大丈夫です。ここでは、あまり詳しく説明するつもりはありません。後日、ビデオドライバーの開発について説明するときに、すべてを詳しく説明します。とりあえず、覚えておいてください。グラフィックスコントローラは、ビデオメモリからのパラレルデータをもとにディスプレイを更新します。これは、ディスプレイの使用時間に応じて自動的に行われます。つまり、ビデオメモリ（デフォルトは0xA0000にマッピングされています）に書き込むことで、現在のモードに応じて、実質的にビデオディスプレイに書き込むことになります。これは、文字を印刷するときに重要です。グラフィックス・コントローラが使用するアドレス範囲を変更することが可能であることを覚えておいてください。初期化の際、BIOSはビデオメモリを0xA0000にマッピングするためにこれを行います。

ビデオモード

「ビデオモード」とは、表示の仕様である。つまり、ビデオメモリがどのように参照され、そのデータがビデオアダプターでどのように表示されるかを記述したものです。

VGAは2種類のモードに対応しています。APAグラフィックス」と「テキスト」です。

APAグラフィックス

APA（All Points Addressable）とは、ビデオモニターやドットマトリックスなど、ピクセルアレイで構成されたデバイスにおいて、すべてのセルを個別に参照できる表示モードのことである。ビデオディスプレイの場合は、すべてのセルが「ピクセル」を表し、すべてのピクセルを直接操作することができます。そのため、ほとんどのグラフィックモードがこの方式を採用しています。このピクセルバッファを変更することで、画面上の個々のピクセルを効果的に変更することができます。

ピクセル

「ピクセル」とは、ディスプレイ上で表現できる最小単位のこと。ディスプレイ上では、色の最小単位を表しています。つまり、基本的には1つのドットである。各ピクセルのサイズは、現在の解像度とビデオモードに大きく依存します。

テキストモード

テキストモードとは、APAのように、画面上のコンテンツを内部的にピクセルではなく文字で表現する表示モードです。テキストモードを採用しているビデオコントローラでは、2つのバッファを使用します。1つは、表示される各文字のピクセルを表すキャラクターマップ、もう1つは、各セルにどのような文字があるかを表すバッファです。文字マップバッファを変更することで、文字そのものを変更することができ、新しい文字セットを作成することができます。また、各セルにどのような文字が入っているかを表すスクリーンバッファを変更することで、画面に表示される文字を変更することができます。テキストモードの中には、文字の色や、点滅、下線、反転、明るくするなどの属性を設定できるものもあります。

MDA, CGA, EGA

VGAはMDA、CGA、EGAをベースにしていることを忘れてはならない。VGAはこれらのアダプタが行うモードの多くをサポートしています。これらのモードを理解することで、VGAの理解が深まります。

MDA

私が生まれる前（真面目な話）の1981年、IBMはPC用の標準的なビデオディスプレイカードを開発した。それが「モノクロディスプレイアダプター（MDA）」と「モノクロディスプレイ・プリンタアダプター（MDPA）」である。

MDAには、グラフィックモードは一切ない。唯一のテキストモード（モード7）では、80列×25行の高解像度のテキスト文字を表示することができた。

このディスプレイアダプターは、古いPCで使われていた一般的な規格でした。

CGA

1981年には、IBMがCGA (Color Graphics Adapter) を開発し、PCの最初のカラーディスプレイ規格となりました。

CGAは、1ピクセルが4バイトに制限されていたため、16色のカラーパレットしかサポートしていなかった。

CGAは、以下の2つのテキストモードと2つのグラフィックモードをサポートしていた。

40x25文字（16色）テキストモード 18x25文字
(16色) テキストモード 320x200ピクセル（4色）
グラフィックモード

640x200ピクセル（モノクロ）グラフィックモード

ディスプレイヤアダプタを使って、「文書化されていない」新しいビデオモードを作成したり、発見したりすることができます。これについては後で詳しく説明します。

EGA

1984年にIBMが発表したEGA (Enhanced Graphics Adapter) は、最大640×350ピクセルの解像度で16色のディスプレイを実現した。

VGAアダプタは、80x86マイクロプロセッサ・ファミリーと同様に下位互換性があることを覚えておいてください。このため、下位互換性を確保するために、BIOSは80列×25行をサポートするモード7 (MDAのオリジナル) で起動します。これは、私たちにとって重要なことです。

ビデオメモリー

メモリーマップドI/O (MMIO)

Memory Mapped I/Oについてご存知の方は、この部分は読み飛ばしてください。

プロセッサは、RAMやROMデバイスからの読み出しで動作することができます。アプリケーション・プログラミングでは、このようなことはまずありません。これを可能にするのがMMIOデバイスです。Memory Mapped I/Oは、ハードウェアデバイスが自身のRAMやROMをプロセッサの物理アドレス空間にマッピングすることを可能にします。これにより、プロセッサは、アドレス空間内のその場所へのポインタを使用するだけで、さまざまな方法でハードウェアRAMやROMにアクセスできるようになります。これは、MMIOデバイスが、プロセッサやシステムメモリが使用するのと同じ物理アドレスとデータバスを使用することで可能になります。

しかし、メモリマップドI/Oは、プロセッサの物理アドレス空間へのマッピングであり、実際のコンピュータメモリではないことを忘れてはならない。アーキテクチャによっては、MMIOデバイスマッピングを使用するか、その後ろに隠れているシステムメモリを使用するかをバンクスイッチで切り替えることができるものもありますが、そうでないものもあります。これが意味するところは、MMIOデバイスによって "隠された" 実際のシステムメモリのアドレスにアクセスできないということです。例えば、CMOS RAMメモリは0x400番地の物理アドレス空間にマッピングされています。これはメインシステムメモリとは異なります。ポインタで0x400にアクセスすると、MMIOに対して常にCMOS RAMメモリにアクセスすることになります。i86アーキテクチャでは、システムメモリのこの場所にアクセスすることはできません。

MMIOデバイスは、限られたシステムメモリで高解像度のビデオ表示を可能にしたり、システムメモリ内では失われてしまう情報を、バッテリーによって最新の状態に保たれたデバイス (CMOS RAM) から得ることを可能にするなど、ハードウェアをよりコントロールすることができます。MMIOデバイスのもう一つの例は、システムBIOS ROMそのものです。MMIOは、システムの物理アドレス空間にマッピングされたROMからプロセッサがBIOSを実行することを可能にします。すごいでしょう？

これがグラフィックと何の関係があるのかと思われるかもしれません。ビデオメモリは、物理アドレス空間にマッピングされたRAMです。ビデオメモリは、MMIOを使用するビデオディスプレイデバイスによって管理されます。MMIOメモリがどのように管理されているかはデバイス次第であり、必ずしも直線的ではありません。グラフィックスモードによって、このメモリの扱い方が異なるため、MMIOデバイスであることを理解することが重要です。

MMIOのアドレス空間領域の面白いところは、ページングにより、任意の仮想アドレスにマッピングし、そのアドレスからアクセスできることです。つまり、例えばビデオメモリを任意の仮想アドレスにマッピングし、その仮想アドレスからビデオメモリにアクセスすることができるのです。これはもちろん、物理アドレス空間のフレームにページがマッピングされることと関係があります。

また、MMIOのメモリはシステムメモリにはないことも覚えておいてください。コンピュータのシステムメモリは、アクセスしようとしているMMIOアドレスのサイズより大きい必要はありません。例えば、システムメモリが2GBしかない場合で

も、0xFC000000の物理アドレス空間にマッピングされたRAMがあれば、MMIOデバイスにエラーなくアクセスすることができます。

この文字が見えますか？私はあなたのコンピュータの中にいて、ビデオRAM（VRAM）に存在しています。VRAMとはビデオメモリのこと、ビデオフレームバッファとも呼ばれています。目の前にあるすべてのピクセルと、それ以上のものが含まれています。

標準VGA

ビデオメモリは、通常はビデオカードやオンボードビデオアダプターなどのビデオデバイスの内部に格納されている。標準的なVGAカードには256KBのVRAMが搭載されている。しかし、SVGA+カードでは、より多くのビデオメモリを搭載しているものも珍しくありません。

しかし。やはり、高解像度のビデオモードでは、すべての画素をなんとか保存しなければなりませんよね。

第7章のメモリマップを覚えていますか？標準VGAのメモリは、0x000A0000～0x000BFFFFにあることがわかります。

$0xBFFFF - 0xA0000 = 0xA0000$ で、655360バイト、つまり640KBです。

ここで覚えておいていただきたいのは、ビデオメモリはPCのアドレス空間ではこの位置にマッピングされているということです。

つまり、ここに書き込むということは、ビデオアダプタにあるビデオメモリに書き込むということです。これはMemory Mapped I/Oの一種です。

ビデオメモリにアクセスするときは、通常、実際のビデオRAMへの「ウィンドウ」を使ってアクセスします。これは典型的なものです。0xA0000 - EGA/VGAグラフィックモード(64KB)

0xB0000 - モノクロテキストモード(32KB)

0xB8000 - カラーテキストモードおよびCGA (32 KB)

モードごとにアドレスのマッピングが異なるため、モノクロディスプレイアダプターとカラーアダプターを同一マシン上で組み合わせることが可能です。これにより、デュアルモニター構成のパソコンでも問題なく動作させることができます。もちろん、これは標準的なVGAです。

スーパーVGA

スーパーVGAやその他のディスプレイアダプタは、一般的に動作が異なります。スーパーVGA以上の解像度のディスプレイアダプタでは、VRAMが高いアドレス範囲にマッピングされていることが珍しくありません。通常は標準VGAのメモリマップ範囲をサポートしていますが、高解像度のビデオモードや追加機能を提供するために、他のメモリ範囲を使用することもできます。例えば、私のNVideo GeForce 7600 GTには、使用可能な4つのメモリ範囲があります。0xA0000～0x000BFFFF（見覚えがありますか？）、0xFC000000～0xFCFFFFFF、0xD0000000～0xDFFFFFFF、0xFD000000～0xFDFFFFFFです。これはあなたのシステムでは異なる可能性があります。

リニアフレームバッファ (LFB)

現在のディスプレイのビデオメモリ全体を物理アドレス空間にマッピングすることができれば、リニアフレームバッファのように動作するように設定することができます。リニアフレームバッファとは、ピクセル単位でパックされたフレームバッファのことです、直線的に読み書きすることができます。例えば、`buffer[0]`はバッファの1番目の要素、`buffer[1]`は2番目の要素で、特に何もありません。まあ、実際にはありますが。標準的なVGAはLFBモードをサポートしていません。上記のモード0x13を覚えていますか？これは、リニアフレームバッファの効果を生み出す唯一の標準VGAビデオモードです。これはちょっとわかりにくいかもしれませんね。結局のところ、ビデオメモリの読み書きが直線的でない場合、どのように「他の」方法で読み書きできるのでしょうか？これは標準VGAがプレナーデバイスであることと関係があります。これについては、次のセクション以降で説明します。

バンク切り替え

スーパーVGAやそれ以上の解像度のビデオモードは、アダプタに搭載されている完全なビデオメモリへの「ウィンドウ」を使用する方法を提供します。例えば、上記のグラフィックモードでは、0xA0000～0xB0000の間の64KBの領域に制限されています。これを「窓」とし、この64Kの窓を「移動」させることができれば、もっと大きなビデオメモリ領域にアクセスすることができます。たとえば、次のようにになります。



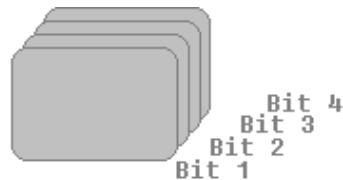
```
unsigned char* vidmem = 0xa0000;

vidmem[0] = 0; //! writes to topleft of screen
movwindow (1); //! moves window (see second picture)
vidmem[0] = 0; //! writes to topleft of screen + size of window
vidmem[0] always points to the topleft of the window, not screen
```

これを「バンク切り替え」といいます。バンク」とは、より大きなビデオメモリへのウィンドウのことです。標準的なVGAは64Kのグラフィック領域しか持たないため、ウィンドウのサイズは通常64Kです。

プランナーメモリー

さて、ここで少し厄介なことが起こります。標準的なVGAのモードはプレーナーメモリモードで動作します。これはVGAのネイティブメモリモデルです。



4 planes, each sharing the same 64K window.
Each pixel's bits share the same location but
different plane.

```
unsigned char* vmem = 0xa0000;

vmem[0] = pixel_bit_0
set_plane(1)
vmem[0] = pixel_bit_1
set_plane(2)
vmem[0] = pixel_bit_2
set_plane(3)
vmem[0] = pixel_bit_3
```

上記は、Mode 12hのプレーナーメモリーフォーマットの例です。モード12hは1ピクセルあたり4ビット。画素を描画するには、プレーナーのビットをセットしたりアンセットしたりする必要がある。これをよく理解するために、64kブロックのビデオメモリがあるとします。それを平らな紙のように想像し、その後ろにさらに3枚の紙を置きます。それぞれの紙は、この同じ64kのメモリ領域を共有する64kの「プレーン」です。それぞれのプレーンは、それが使用されているピクセルに関する少しの情報を保持しています。

この章では必要ないので、プレーナーのメモリとその仕組みを理解することを心配する必要はありません。しかし、VGAとモード12hについて詳しく説明する際には重要になります。本章では、プレーナーメモリの動作の詳細を隠したモード0x13を使用するため、今は必要ありません。

奇数/偶数メモリのアドレッシング

Odd / Even Memory Addressingは、Planer Memoryモデルを使用し、すべてのテキストモードで使用されます。すべての偶数アドレスはプレーン0または2で動作し、奇数アドレスはプレーン1または3で動作します。例えば、以下のようになります。

Memory Address	Plane	Offset in plane
0	Plane 0	Offset 0
1	Plane 1	Offset 0
2	Plane 0	Offset 2
3	Plane 1	Offset 2

テキストモードでのビデオメモリへの書き込みを思い出してください。

```
unsigned char* vmem = 0xb8000;
vmem[0] = 'a'; // plane 0 [character plane] offset 0
vmem[1] = 0x7; // plane 1 [attribute plane] offset 0
vmem[2] = 'b'; // plane 0 [character plane] offset 2
vmem[3] = 0x7; // plane 1 [attribute plane] offset 2
```

テキストモードでは、プレーン0に文字コードが格納され、プレーン1に属性バイトが格納されます。プレーン2には、フォントデータが格納されます。ビデオメモリへの書き込み時にプレーン2を上書きすると、BIOSが起動時にインストールしたフォントを上書きすることになる。つまり、グラフィックモードでプレーン2を上書きして、テキストモードに戻ると、フォントデータが壊れているため、BIOSのテキスト出力ルーチンが期待通りに動作しないということです。

テキストモードに戻したい場合は、独自のフォントを保存するか、デフォルトフォントをバックアップしてプレーン2に書き戻してからテキスト出力ルーチンを使用する必要があります。

本章ではプレーナーメモリーモデルを使用しないため、Odd/Evenアドレッシングモデルも使用しません。

カラーパレット

パレットは、ルックアップテーブルのようなものです。カラーパレットは、色のルックアップテーブルのようなものです。例えば、実際の色情報のリストをテーブルに格納しておきます。そして、そのテーブルへのインデックスを別のテーブルすることができます。

Index Table | Color Palette

0	red(0), green(0), blue(0)
1	red(0), green(0), blue(1)
2	red(0), green(1), blue(0)
...	

上の例では、インデックスを使用するだけで、好きな色を参照できます。ルックアップテーブル（カラーパレット）が作成された後は、色を参照したいときにはいつでもインデックスを使うことができるので、ストレージスペースを大幅に節約できます。

例えば、カラーパレットを使用するビデオモードでは、ビデオメモリーがインデックスバッファの役割を果たします。そのため、先ほど作成したパレットを使ってピクセルを描画するには、使用したい色のインデックスを書き込めばよい。

```
unsigned char* p = 0xa0000;
p[0] = 0; // black pixel
p[1] = 1; // blue pixel
p[2] = 2; // green pixel
```

VGAでは、カラーパレットはハードウェアで処理されます。パレットの色を好きなようにコントロール、変更することができます。ただし、パレットの操作にはVGAハードウェアのプログラミングが必要なので、ここではありません。VGAハードウェアの話になった時に説明しますので、ご安心ください。

パレット・アニメーション

さて、ここで一步引いて考えてみましょう。上の例をもう一度見てみましょう。ビデオディスプレイでは、ピクセルの色をインデックスで決定します。例えば、カラーパレットのインデックス1（上の例ではLike）が別の色に変わったとしたらどうでしょう？上の例では、カラーパレットのインデックス1は明るい青色です。つまり、パレットビデオモードであれば、ビデオメモリのどこかに「1」を書き込むと、いつでもその鮮やかな青色になるということです。つまり、memset(vidmem, 1, VIDMEM_SIZE) を実行するだけで、ビデオディスプレイをこの色にすることができます。かっこいいでしょう？

ビデオディスプレイの表示色は、「カラーパレット」テーブルの中にあるインデックスの色で決まることがわかっているので、どのパレットエントリの色でも変更することができます。これにより、パレットの色を何らかの方法で更新するだけで、画面上の色を変化させることができます。これを「パレットアニメーション」と呼びます。

パレットアニメーションは、炎のアニメーションや氷のエフェクトなど、見栄えのするクールなエフェクトを多数作成することができます。

モード 0x13

アブストラクト

ビデオモード0x13は、256色320x200の解像度を持つIBMの標準的なVGA BIOSのモード番号である。256色のパレットを使用し、正方形のピクセルを持たず、パックドピクセルフレームバッファとしてビデオメモリにアクセスすることができた。これはどういうことかというと、あたかもリニアバッファであるかのようにビデオメモリへのアクセスを可能にしたのである。ビデオメモリへのポインタを取得するだけで、ポインタがunsigned char*であると仮定して、pointer[0]=ピクセル1、pointer[1]=ピクセル2...となります。これは、特定のハードウェアレジスタの設定（ビデオモードの「設定」）によって可能になります。標準的なVGAは、それ自体ではこのようなビデオメモリへのアクセスを提供しません。

ここで重要なのは、ビデオモードでは、解像度、ビデオメモリへのアクセス方法、そのモードを動作させるためのハードウェアの設定などが定義されていることです。ここですべてを理解できなくても気にしないでください。後の章でVGAハードウェアについて説明する際に詳しく説明します。

ビデオモード0x13は簡単に扱える（そして速い）ので、この章ではこれを使うことにしました。他のモードの中には、VGAハードウェアの経験が必要なものもありますが、その複雑さゆえに、この章では避けたいと思います。しかし、心配しないでください。後ほど、いくつかのモード（640x480x4カラーのモード12hなど）を紹介する予定です。

ビデオモード0x13は、DOS時代にビデオゲームによく使われていましたが、プログラムが簡単で速度も速いのが特徴です。

幅320×高さ200ピクセルの解像度で、256色のカラーパレットを持つビデオ構成です。プレナー方式のビデオメモリモードですが、リニアフレームバッファ(LFB)として動作するため、プログラムが簡単です。

カラーパレット

モード0x13は、256色のカラーパレットを持つ。モード0x13のビデオメモリには、パレットインデックスが格納されているだけで、ビデオデバイスはインストールされているパレットのカラーテーブルからレンダリングする色を決定する。デフォルトでは、このようなカラーテーブルになっています。



モード 13h カラーパレット

例えば、上の図を見ると、最初の色(0)は黒、色1は青、色2は緑、などとなっています。これらの色をビデオディスプレイに書き込むには、上記のルックアップテーブルでこれらのインデックスを使用します。

```
unsigned char* p = 0xa0000;
*p = 0; //black pixel
*(p++) = 1; // blue pixel
*(p++) = 4; // red pixel
*(p++) = 255; //white pixel
```

上のコードを上の表と比較して、インデックスがパレットの色と一致していることを確認してください。

パレットの変更

パレットを好きな色に変更することは可能です。しかし、そのための簡単なBIOS割り込みはありません（とにかくVBEを使わないと無理です）。ほとんどの割り込み呼び出しは、VGA Digital to Analog Converter (DAC) の内部にある個々のパレットレジスタまたはすべてのパレットレジスタを設定または取得するために使用されます。これにはVGAハードウェアの知識が必要ですが、この章では簡単にするために避けたいと思っています（心配しないでください、近いうちにそれをカバーする予定です！）。

ビデオBIOSインターフェース

VGAビデオBIOSインターフェイスは、ビデオ割り込みのセットです（ソフトウェア割り込み0x10）。これらはBIOS割り込みであるため、リアルモードまたはv86モードでしか使用できません。

ビデオモードの設定

INT 0x10 機能 0

ビデオモードの設定は、BIOS割り込み0x10の機能0：入力を呼び出すことで行うことができます。

AH = 0

AL = ビデオモード出力

AL = ビデオモードフラグ (Phoenix, AMI BIOS)

AL = CRTコントローラ(CRTC)モードバイト (Phoenix 386 BIOS v1.10)

CRTCは、ビデオハードウェアを直接プログラムする場合に必要となるコントローラの一つであるため、今後多くの場面で目にすることになるでしょう。

この割り込みでは、任意のテキストモードやビデオモードを設定することができます。例えば、以下は320x200x8ビット[モード0x13]に切り替えます：（全てのコードサンプルはデモにあることを覚えておいてください）。

```
mode13h:  
    mov ah, 0  
    mov al, 0x13  
    int 0x10  
    ret
```

簡単でしょう？

以上のことから、グラフィックモードにすることができます。確かに、リアルモードかv86モードでしか動作しませんが、これほど簡単な方法はありません。ビデオモードを理解していなくても心配ありません。ビデオモードについては後で説明します。

ビデオモードの取得

INT 0x10 機能 0xF

BIOS割り込み0x10関数0xF：入力を呼び出すことで、ビデオモードを取得することができます。

AH = 0xF

出力

AH = 文字列数 AL = 表示モード番号

BH = アクティブページ

この割り込みは簡単なもので、現在のビデオやテキストのモードを取得するのに使用できます。まだ「アクティブなページ」の部分は気にしないでください。ビデオモードを理解していなくても心配ありません。後ほど説明します。

```
getMode:
    mov ah, 0xf
    int 0x10
    ret
```

その他のビデオBIOSインターラプト

INT 0x10 機能 0xB/BH=1

ビデオBIOSのINT 0x10の機能0xB：入力を呼び出すことで、パレットの設定が可能です。

AH = 0xB

BH = 1

BL = パレットID

00h 背景、緑、赤、茶/黄 01h 背景、シアン、マゼンタ、白

この割り込みは、すべてのシステムでサポートされているわけではありません。

INT 0x10 機能 0xC

この割り込みを使って、ディスプレイにピクセルを書き込むことができます。

入力

AH = 0xC

BH = ページ番号 AL = ピクセルカラ

ー

ビット7がセットされている場合、256色モードを除き、値はスクリーン上でXORされる CX = 列

DX=ロウ・アウトプット

AL = ピクセルの色

この割り込みは、グラフィックモードでのみ使用できます。

INT 0x10 機能 0xD

Video BIOS INT 0x10関数0xB: Inputを呼び出すことで、ピクセルを読み取ることができます。

AH = 0xC

BH = ページ番号 CX =

コラム

DX = ロウ

この割り込みは、グラフィックモードでのみ動作します。

プリミティブ

最初のピクセルをプロットする

"あらゆるビデオゲームを作る秘訣は、ピクセルの色を変える能力である。"- Teej

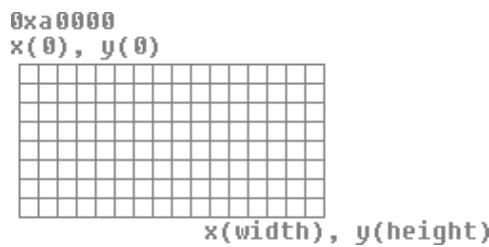
この章では多くのことを学びましたが、まだ画面上にピクセルを描くことができません。どうなっているのでしょうか？この章の最後に、グラフィックスの最も基本的なプリミティブである「ピクセルのスクリーンへの描画」の基本を紹介することにしました。

今回はMode 0x13での作業なので、リニアフレームバッファのように動作することを覚えておいてください。つまり、vidmem[0]がビデオメモリの1バイト目、vidmem[1]が2バイト目ということになります。また、モード0x13では、各ピクセルの1バイトをカラーパレットへのインデックスとして使用することを覚えておいてください。つまり、次のようにピクセルを簡単に書くことができるのです。

```
unsigned char* p = 0xa0000;
    p[0] = 1; // blue pixel
```

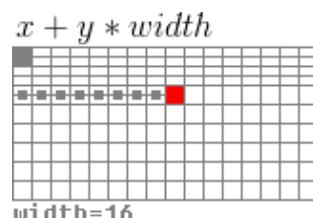
かっこいいでしょう？これだけで、ピクセルが出来上がりました。

直交座標系で考えるとわかりやすいですね。このシステムでは、XやYといった座標を使って、このような2次元のグラフ上の位置を表現します。



ビデオメモリの左上隅は $v = [0,0]$ で、 v は2次元のベクトルである。これはビデオディスプレイの最初のピクセルです。最後のバイトは、 $v = [width, height]$ です。各座標がピクセルであると仮定すると、画面上の任意の位置にピクセルを描くことができる公式を思いつくことができます。

上のグラフで $v = [0,0]$ の位置からスタートしたとします。この位置に幅を加えると、常に元の位置のすぐ下になってしまいます。例えば、上のグラフでは、幅=16です。左上からスタートしたとすると、右に16を数えると、スタートした場所の真下（次の線上）にいることになります。このことから、 $y * width$ とすることで y を計算することができます。その後、 x （その行のオフセット）を加えれば、式が完成します。



任意の $[x,y]$ の位置にピクセルを描画するには、 $x + y * width$ という式を使います。これで、次のような簡単なルーチンを作ることができます。

```
;-----;
; renders pixel
; cl = color ax = y bx = x
; es:bp = buffer
;-----;
pixel:
; [x + y * width] = col

    pusha
    mov di, VGA_MODE13_WIDTH
    mul di ; ax = y * width
    add ax, bx ; add x
    mov di, ax
    mov byte [es:bp + di], cl ; plot pixel
    popa
    ret
```

es:bpはビデオディスプレイ、またはレンダリングしたい他のバッファを指します。clは使用したいカラーインデックス、axはY位置、bxはX位置です。

画面の消去

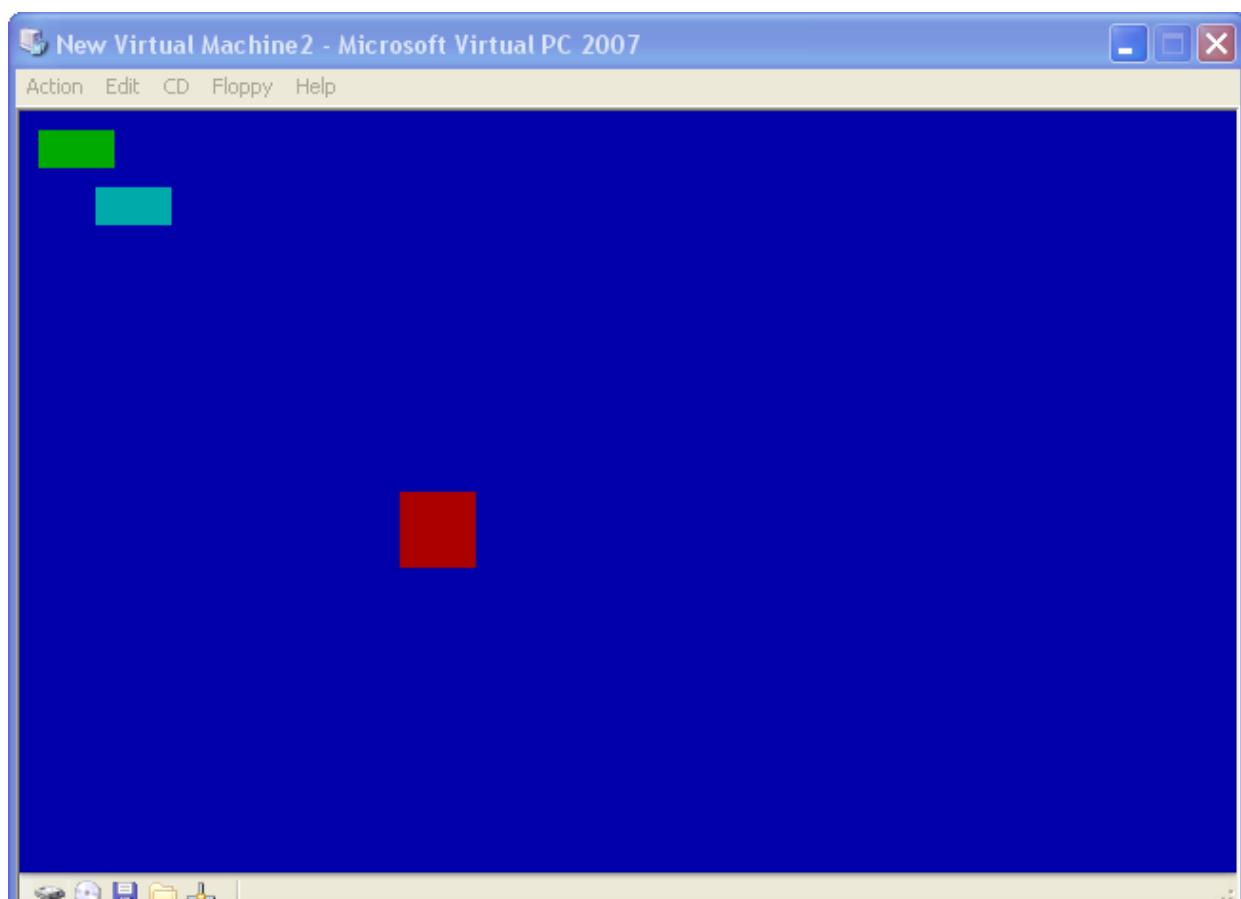
画面を消去するにはいくつかの方法があります。これは重要なことで、ビデオモードを切り替えたときに、画面にゴミがたくさん表示されることがあります。

1つの方法としては、上記のピクセルルーチンをwidth * heightの回数だけ呼び出すことができます。もっと良い方法は、一度に複数のピクセルを書き込むことです。例えば、モード13hのピクセルサイズが1バイトであることを知っていれば、ワードサイズレジスタに2バイト（2ピクセル）を格納して、それを利用することができる簡単な例を示します。

```
;-----;
;  clear screen
;  cl = color
;-----;
clrscr:
    pusha
    mov dl, cl ; dx = 2 pixels
    mov dh, cl
    mov cx, 0
    xor di, di
    .l:
    mov word [es:bp + di], dx ; plot 2 pixels
    inc di ; go forward 2 bytes
    inc di
    inc cx
    cmp cx, (VGA_MODE13_WIDTH * VGA_MODE13_HEIGHT) / 2 ;end of display?
    jl .l
    popa
    ret
```

es:bpはビデオメモリまたは他のバッファを指し、clは使用したい色を指します。

Demo



デモダウンロード

このデモでは、これまで説明してきた内容に、水平方向の線を描画する「line」というルーチンを追加して、少しだけスペースを加えています。

結論

この章では以上です。

次の章では、VESA VBEと、それを使って高解像度のグラフィックモードをどのように扱うかについて説明します。また、スーパーVGA、バンクスイッチング、ダブルバッファリング、トリプルバッファリング、ページフリッピングなどのグラフィックコンセプトについても説明します。そうです、VBEで高解像度を実現するのです：)

また、次の章ではC言語に戻り、グラフィックプリミティブをいくつか取り上げる予定です。VGAのハードウェアについては、VBEの後に取り上げようと思っています。VGAのハードウェアは非常に複雑なので、グラフィックやVGAの複雑な話題はもう少し後にしたいと思っています。

次の機会まで。

~マイク

BrokenThorn Entertainment社。現在、DoEとNeptune Operating Systemを開発中です。質問やコメントはありますか？お気軽にお問い合わせください。

あなたも記事の改善に貢献したいと思いませんか？もしそうなら、ぜひ私に教えてください。

[Home](#)





オペレーティングシステム開発シリーズ

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

"複雑さをコントロールすることは、コンピュータ・プログラミングの本質である" - Brian W. Kernighan

グラフィックス2: VGAおよびSuperVGA

by Mike, 2013

1. Introduction

Welcome!

1. 前回の記事では、Video BIOSファームウェア上のVGAに重点を置いて、ビデオデバイスに関する概念を紹介しました。今回は、VGA、SuperVGA、Video BIOSをサポートするビデオデバイスのハードウェアとファームウェアのインターフェースの違いを見て、それらのインターフェースをサポートする共通のビデオインターフェースをC言語で実装してみます。以下は、そのトピックです。
 2. VGAハードウェア
 3. VGA BIOS
 4. Vesa BIOS Extensions
 5. Bochs VBE Interface

まず最初に、パーソナルコンピュータで最も多くサポートされており、最も古い規格の1つであるVGA (Video Graphics Array) 規格を紹介する。VGA規格は、標準的な方法でビデオハードウェアと対話する方法を提供していますが、低解像度の表示モードや、最新のディスプレイメモリに存在するグラフィックアクセラレーションのサポートがないなどの制限があります。ここでは、ハードウェアインターフェースとVGA BIOSファームウェアインターフェースの両方を見ていきます。VGA BIOSインターフェースは、VGAハードウェアインターフェースよりもはるかにシンプルですが、realまたはv86プロセッサモードでしか使用できません。ハードウェアインターフェースは、どのプロセッサモードでも使用できます。続いて、VESA BIOS Extensions (VBE) に移ります。VBEは、VESA (Video Electronic Standards Association) が策定した規格で、高解像度の表示モードやモニター機能をサポートするためのBIOS拡張機能の標準セットを提供しています。BIOSの拡張機能であるため、すべてのパソコンがサポートしているわけではありません。また、リアルモードまたはv86モードからしか使用できません。

この章の目的は、ビデオモードの変更とディスプレイメモリへのアクセスを行うことです。この章の終わりには、VGAまたはSuperVGAを使用したデモを作成することができるでしょう。この後の章では、グラフィックスに焦点を当てていきます (SuperVGAのハードウェアグラフィックスをサポートする場合もあります)。

1.1. BIOSとの連携

本章のメイントピックに入る前に、少し回り道をして、BIOSを詳しく見てみましょう。前回の記事で、VGA用のBIOSサービスのいくつかを使用したこと思い出してください。これらのサービスを利用する場合、ソフトウェアはリアルモードまたはv8086モードで動作しなければなりません。これはプロテクトモードやロングモードのソフトウェアにとっては問題となります。そのため、先に進む前にこれを解決する方法を見つける必要があります。ただし、BIOSサービスを使用する予定がない場合は、このセクションをスキップしてください。

1. プロテクトモード (ロングモードではない) からBIOSを呼び出すには2つのアプローチがあります。
 2. リアルモードに落としてBIOSを呼び出す
 3. Virtual 8086 mode

第一の方法は、よりシンプルですが、より完全に設計されたシステムでは非常に不便です。2番目の方法は、最も一般的に使用されている方法ですが、最も困難な方法もあります。ユーザー モード、割り込みディスパッチ、タスクスイッチ、命令エミュレーションが必要です。

方法1

1番目の方法は、ソフトウェアが必要に応じてプロテクトモードからリアルモードに切り替えられるようにすることです。この方法は、ソフトウェアの設計に一定の制限 (仮想メモリをサポートしないとか、上位半分のカーネルをサポートするとか) を加えないと、サポートが複雑になり、その価値がなくなってしまいます。しかし、この方式は最もシンプルな実装方法であり、追加のソフトウェアサポートも最小限で済みます。以上の理由から、今回のデモではこの方法を採用しましたが、ソフトウェアシステムが十分に大きい場合には、方法2を使用することを強くお勧めします。

この方法を実装するためには、32ビットのプロテクトモードと16ビットのリアルモードの間のインターフェイスとして機能するルーチンまたはルーチンのセットが必要です。これらのルーチンは、ルーチンの入力値と出力値を保持しながら、以下のことを行う必要があります。

1. 予約しなければならない現在のシステム状態を保存します。最も基本的なケースでは、プロテクトモードのスタックとIDTRです。
2. Disable hardware interrupts (CLI instruction).

3. Reload original IVT. This is done by setting IDTR.size to 0xffff and IDTR.base to 0.
4. Perform a jump to 16 bit protected mode.
5. Disable protected mode by clearing CR0.PM bit
6. Perform a jump to 16 bit real mode code.
7. Set all real mode segments, enable interrupts and call BIOS.
8. Perform a jump to 32 bit protected mode.

9. Restore saved system state. In the most basic case as in (1) this is the selectors, protected mode stack, and IDTR.

このルーチンは、システムがサポートするものがより要求されるようになると、非常に複雑になります。上記のリストは大変そうですが、システムがページングを使用しておらず、カーネルイメージが1MB以下であれば、難しいことはありません。つまり、プロジェクトのベースアドレスが64Kで、ページングが禁止されていると仮定することで、ルーチンを比較的シンプルに保つことができます。

デモでは、BIOSの呼び出しにio_servicesというメソッドが使われています。上記の手順でリアルモードに落ちます。io_servicesは以下のようになります。

```
extern void io_services (unsigned int num, INTR* in, INTR* out);
```

numは割り込み番号、inはINTR構造体へのポインタ、outは出力INTR構造体へのポインタです。INTRは、レジスタの値を格納する構造体のセットです。io_servicesとINTRはどちらもかなり大きいので、本文では省略します。デモではbios.asmを参照してください。

ソフトウェアがロングモードの場合は、デバイスを直接プログラムするか、エミュレーターを書くしかありません。

方法2

1. 2つ目の方法は、v8086モードを使用することです。この方法は、BIOSファームウェアの呼び出しをサポートするための長期的な方法ですが、最も要求の厳しい方法もあります。最低でもv8086モードでは、OSが以下をサポートしている必要があります。

2. ユーザーモードのプロセス
3. Task switching
4. Interrupt dispatching
5. Instruction emulation

仮想8086モードは、ユーザーモードプロセスとしてのみ実行できます。しかし、これには問題があります。ユーザーモードのプロセスはカーネルモードの命令(intなど)を実行できないため、BIOSを呼び出すことができず、目的が達成できません。つまり、v8086プロセスがint(割り込み)命令を実行すると、GPF(General Protection Fault)が発生してしまうのです。これを解決するにはどうしたらいいでしょうか？

1. ここで完全に解決策がないわけではありません。v8086プロセスはBIOSを呼び出すことはできませんが、カーネルは呼び出すことができます。

GPFが発生すると、事実上、カーネルが呼び出されます。カーネルのGPFハンドラは、GPFが発生した原因を検出し、次のように対処します。

2. 現在のプロセスを確認する v8086 フラグ
3. If set; call v86_monitor

設定されていない場合は、GPFを続行し、場合によってはプロセスを終了させる

v86_monitorは、すべてのv8086プロセスのカーネルGPFハンドラから呼び出される特別な関数です。ここでは、v8086プロセスがBIOSを呼び出そうとする（または、カーネルモードの命令を使おうとする）たびに、カーネルGPFハンドラが呼び出され、v8086タスクを「監視」するためにv86_monitorが呼び出されるということです。

v86_monitorは、v8086タスクが使おうとした問題命令をエミュレートする役割を持つv8086モニターを実装します。例えば、v8086モニターは、問題のある命令（CPUから与えられたCS:EIPを持つ）を割り込み呼び出しとして検出し、IVT [n] (n = 呼び出すBIOS番号) を呼び出すことで、それをエミュレートします（ただし、IVT命令のポインタは、線形ではなく、segment:offset形式であることを思い出してください）。

2. Video Graphics Array (VGA)

SuperVGAだけを見たい方は、この項を読み飛ばしても構いません。

VGA (Video Graphics Array) は、1987年にIBMのPS/2コンピュータ用に初めて導入されたディスプレイハードウェアのデザインであるが[1]、現在ではディスプレイの標準規格として各組織で広く採用されている。サポートされているビデオモードの最高解像度は640x480x16色です。PCメーカーに広く採用されたため、VGAは現在のPCでもサポートされている最も古い規格の1つとなっています。

VGAについてIBMのXGA (Extended Graphics Array) 規格が導入されたが、メーカーごとに拡張機能が実装され、最近のPCではSuperVGAアダプタが一般的に使われている。ほとんどのSuperVGAカードは、VGA規格との下位互換性がある。

この記事は、VGAの複雑さゆえに、網羅的な説明にはならないことをご了承ください。VGAハードウェアの詳細については資料[2]と[3]を参照してください。また、VGAに関する多くの大規模な書籍を読むことをお勧めします。

2.1. Video modes

VGAがサポートするビデオモードとモード番号には、標準的なセットがあります。ビデオモードとは、ディスプレイの構成と、解像度（1ピクセルあたりのビット数）、色数、メモリモードなどのプロパティのことを指します。標準的なビデオモード番号は、0h、1h、2h、3h、4h、5h、7h、Dh、Eh、Fh、10h、11h、12h、13hです。モード番号自体は特別なものではなく、ビデオBIOSが特定のビデオモードを参照するために使用するものです。もっと多くのモードがあるかもしれません、それらは標準ではありません。

0h	40x25 Text	16 Color	Dh	320x200	16 Color
1h	40x25 Text	16 Color	Eh	640x200	16 Color
2h	80x25 Text	16 Color	Fh	640x350	2 Color
3h	80x25 Text	16 Color	10h	640x350	16 Color
4h	320x200	4 Color	11h	640x480	2 Color
5h	320x200	4 Gray	12h	640x480	16 Color
7h	80x25 Text	2 Color	13h	320x200	256 Color

標準のVGAでサポートされている最高の解像度は、640x480x16色のモード12hです（面白いことに、Windows XPのロゴ画面はモード12hで動作します）。(それ以上の解像度を得るには、後述する「SuperVGA」を使用する必要がある)。

VGAファームウェア

まず、VGAのファームウェアインターフェースと、Video BIOSが提供するファシリティについて見ていきます。前回は、ビデオモードを設定するための割込み0x10関数0を中心に、いくつかの機能を紹介しました。Video BIOSは、より抽象的なインターフェースを通じて、VGAハードウェアの設定、取得、および操作のためのサービスを提供します。おそらく、ソフトウェアがハードウェアを直接制御するよりも、ビデオBIOSの機能を使用する方がはるかに安全で、シンプルで、移植性が高いでしょう。

ここでは、ソフトウェアがビデオサービスに使用できる一般的な設備を紹介します。もちろん、これらはBIOS割り込みなので、リアルモードまたはv8086モードで、BIOSファームウェアを持つシステムでしか使用できません。また、これらはカーネルモードのソフトウェアでしか使用できません。

INT 0x10 Function 0 - Set Video Mode

入力します。

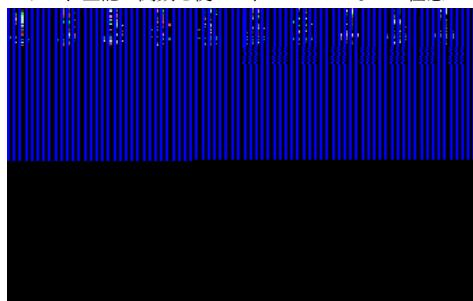
- AH=0
- AL = ビデオモード

出力します。

AL=ビデオモード フラグ (Phoenix, AMI BIOS)

AL = CRTコントローラ(CRTC)モードバイト (Phoenix 386 BIOS v1.10)

これで、上記の関数を使って、Mode 13hなどの任意のVGA BIOSビデオモードを設定することができます。結果は次のようにになります。



VGAモード設定後の結果。

1. 上の画像を見ていたければわかるように、ディスプレイにはたくさんのゴミがあります。この「ゴミ」は、実際にはモード切り替え前にVGAメモリにあったものです。これには、プレーン1のテクスチャ文字やプレーン2のVGAフォントなど、4つのプレーン（これらについては後述します）すべてのものが含まれています。ディスプレイをクリアすると、これらすべてがクリアされる。これは、ソフトウェアがテキストモードに戻らなければならない場合、ディスプレイをクリアすることによってソフトウェアがVGAフォントをクリアしてしまうため、問題となる。この問題を解決するには2つの方法があります。

2. Upload a new copy of the VGA font from the BIOS

3. Save the VGA font before clearing memory and restoring it later.

上記のいずれかを実行することで、グラフィックモードに切り替わり、エラーなくテキストモードに戻ることができます。

INT 0x10 Function B - Set Palette (テキストモードのみ)

入力します。

AH=B
BH=1

BL=パレットID。
00h 背景、緑、赤、茶/黄 01h 背景、シアン、マゼンタ、
白

◆ Example. The following code sets a VGA palette.

```
void vga_set_palette (int id) {
    /* call BIOS */
    INTR in, out;
    in.eax.val = 0xB;
    in.ebx.val = 0x0100 | id;
    io_services (0x10, &in, &out);
}
```

INT 0x10 Function C - Write pixel

入力します。

- ◆ AH=C
BH=ページ番号 AL=ピクセルカラーバイト
CX=列
DX=行

出力します。

AL=ピクセルカラー

◆ Example. The following writes a pixel. This might be omitted in the demos as it shouldn't be used; we only provide it for completeness.

```
void vga_plot_pixel (int col, int x, int y) {
    /* call BIOS */
    INTR in, out;
    in.eax.val = 0x0C00 | col;
    in.ebx.val = 0;
    in.ecx.val = x;
    in.edx.val = y;
    io_services (0x10, &in, &out);
}
```

なお、Bochsエミュレータの一部のバージョンでは、この割り込みをサポートしていません。しかし、VirtualPCには実装されています。上記の割り込みを使って、ピクセルをプロットすることができます。

INT 0x10 Function F - Get Video Mode

入力します。

- ◆ AH=F

出力します。

AH=文字列の数 AL=表示ページ番号
BH=アクティブページ

◆ Example. The following code returns the mode information.

```
void vga_get_mode (unsigned int* col, unsigned int* dispPage, unsigned int* actPage)
{
    INTR in, out;
    /* sanity check */
    if (!col || !dispPage || !actPage)
        return;

    in.eax.val = 0xf;
    io_services (0x10, &in, &out);
    *dispPage = out.ax.r.al;
    *actPage = out.bx.r.bh;
}
```

2.2. VGA Hardware Interface

VGAのハードウェアインターフェイスは非常に複雑で、5つのコントローラと、ポートI/Oアドレス空間からアクセス可能な100以上のハードウェアレジスタで構成されている。現在の標準的なVGAのビデオモードは、標準的なレジスタの構成によって定義されている。その理由は

ポートI/Oアドレス空間を使用することで、ソフトウェアはどのプロセッサモードでもVGAハードウェアとインターフェイスすることができる。もちろん、in/outファミリーの命令が実際にハードウェアとやりとりするためには、ソフトウェアがスーパーバイザレベル（リング0）である必要がある。従って、VGAハードウェアにアクセスできるのはカーネルモードのソフトウェアだけである。ユーザーモードのソフトウェアがアクセスしようとすると、in/out命令の実行時に一般保護フォルトが発生します。

ハードウェアを見る前に、警告があります。この警告は、無効なデータやデバイスがサポートしている範囲外のデータ（例えば、デバイスが安全に動作できない過剰な周波数設定など）に対する保護機能が欠如している可能性があるすべてのディスプレイモニターやビデオカードに適用されます。最近のモニターでは、無効な設定にはエラーメッセージが表示されたり、出力されなかったりします。必ず最初にエミュレーターや仮想環境でドライバーソフトをテストし、ドライバーソフトが正常に動作することを確認してから実際のハードウェアでテストしてください。

ビデオメモリー

1. VGAのメモリレイアウトは、現在アクティブになっているビデオモードの種類によって異なります。VGAは以下のメモリレイアウトに対応しています。
 2. リニア
 3. Planar
 4. Palette
 5. 4 Color mode

次に、これらのモードのすべてと、どのような場合に使用されるかを見てみましょう。メモリモデルの違いを知ることは、ディスプレイのメモリを正しく読み書きするために重要です。ここでは、最も複雑なリニアモードを最後に取り上げます。その前に、まずビデオメモリへのアクセス方法を知っておく必要があります。

システムアーキテクチャの章で見たメモリマップを思い出してください。ここでは、再びそれを紹介します。

```

0x00000000 - 0x000003FF - リアルモードインタラプトベクターテーブル
0x00000400 - 0x000004FF - BIOSデータエリア
0x00000500 - 0x00007BFF - 未使用
0x00007C00 - 0x00007DFF - 当社ブートローダ
0x00007E00 - 0x0009FFFF - 未使用
0x000A0000 - 0x000BFFFF - ビデオRAM (VRAM) メモリ
0x000B0000 - 0x000B7777 - モノクロ・ビデオメモリ 0x000B8000 -
0x000BFFFF - カラー・ビデオメモリ 0x000C0000 - 0x000C7FFF - ビデオ
ROM BIOS
0x000C8000 - 0x000EFFFF - BIOSシャドウ領域
0x000F0000 - 0x000FFFFFF - システムBIOS

```

メモリマップによると、ビデオメモリは0xA0000～0xBFFFFの物理アドレス空間にマッピングされています。これは0x1FFFFFバイトのメモリ、つまり131072バイトで128Kとなります。[2]によると、VGAハードウェアは最大256Kのメモリを持っていますが、マッピングされているのはそのうちの128Kのみです。マップされていないメモリは、VGAのアドレスデコードメカニズムを変更することでアクセス可能になります（ここでは説明しませんが）。以下のいくつかの例でこれを示します。

プラーナーメモリーモード（16ビットカラーモード）

VGAのメモリは、4つのプレーンがそれぞれ64kのメモリとして参照されます。これらは、互いに接続された異なるメモリバンクと考えることができます（実際にはそうではないかもしれません）。これらのウィンドウは、プレーンです。

16色モードでは、1色につき4ビットが使用されます。この4ビットは、各プレーンの同じ位置に格納されています。プラーナーメモリモードでの画素の書き込み例は、レジスタやビデオモードの設定方法を確認した後、少し後に紹介します。残念ながら、ピクセルの書き込みには、書き込み先のプレーンを選択するためのハードウェアレジスタへの書き込みも必要となるため、まだ例を示すことができません。

パレット・メモリー・モード（256色モード）

パレットメモリーモードでは、各画素はカラーテーブルへのインデックスという数字で表されます。このカラーテーブルがパレットである。代表的なのは256色モードで、各画素が8ビットである。その他のパレットモードとしては、16色（4ビットピクセル）や2色のみのモノクロ（1ビットピクセル）モードがある。

カラーテーブルは以下のようになっています。これは16色モードで使われるパレットで、実際、システム起動時のデフォルトはVGAテキストモードのパレットである。

Index	Color name	Index	Color name
0	Black	8	Dark Gray
1	Blue	9	Light Blue

2	Green	10	Light Green
3	Cyan	11	Light Cyan
4	Red	12	Light Red
5	Magenta	13	Light Magenta
6	Brown	14	Yellow
7	Light Gray	15	White

ここでは、テキストモードでのパレットの使用例を紹介します。

DOSのビデオゲームでは256色モードが広く使われていた。このモードは、標準的なVideo BIOSのモード番号にちなんで「モード13h」と呼ばれ、有名になった。モード13hは、非常に高速で、リニアで、一度に256色を画面に表示できるという興味深いモードである。つまり、このモードのビデオメモリは、各ピクセルが1バイトで、メモリ内で隣り合って保存されるリニアなものなのだ。モード13hが面白いのは、VGAがリニアではなく平面的な表示デバイスだからだ。

◆ Example. The following C code defines a function, **pixel_256** that plots a pixel in 256 color mode at some x and y location in mode 13h. Recall that mode 13h is 320x200 and so the pitch (here pitch is width) is 320. Also unlike the VGA BIOS interrupt plot pixel service, this one will work in any PC emulator or virtual machine.

```
#define VGA_VRAM 0xA0000
#define PITCH 320
void pixel_256 (unsigned char color, unsigned int x, unsigned int y)
{
    unsigned char* fb      = (unsigned char*) VGA_VRAM;
    unsigned int offset   = y * PITCH + x;
    fb [offset] = color;
}
```

上記の例は、Mode 13hを使った作業のシンプルさを示すものでもあります。複数のピクセルに連続して書き込むだけで、複数のピクセルをレンダリングできます。256色モードでは、各ピクセルがバイトとして表現されるため、符号なしのcharを使用します。

パレットメモリーモードでは、一度に表示できる色数に上限があります。これは、そのモードで使用されているカラー テーブルのエントリ数と同じです。例えば、256色モードでは、最大で256色までしか画面に表示できません。16色モードでは16色しか表示できない。ソフトウェアでは、カラー テーブル自体を変更して異なる色を表示することもできる。

リニアメモリーモード

リニアメモリモードとは、リニアフレームバッファ (LFB) をサポートするビデオモードのことである。リニアメモリーとは、C言語の配列のように、連続したバイトの配列のことと、モード13hはリニアメモリーモデルを持つモードの一例である。また、悪名高い「モードX」(360×480×256色) や「モードQ」(チーン4 256×256×256色) は、「モード13h」を改良したものだ。

VGAはリニアなデバイスではなく、リニアなメモリモードをサポートしていないことを思い出してほしい。モード13h、モードX、モードQなどのモードは、いずれもリニアメモリモデルのように錯覚させる方法でハードウェアを構成するプラナーモードである。

リニアメモリーモードでピクセルを書き込むには、ピクセル位置のオフセットを計算して、フレームバッファーにピクセルを書き込むだけよい。

ハードウェア

1. VGAのハードウェアデザインは、以下のコンポーネントで構成されています。これらの中には見覚えのあるものもあるかもしれません。
 2. CRT Controller
 3. Sequencer
 4. Graphics Controller
 5. RAMDAC
 6. Video memory
 7. Attribute Controller

すべてのハードウェアレジスターは、I/Oポート空間にマッピングされています。つまり、CPUのin/out命令群を使ってアクセスすることができるのです。ほとんどのコントローラは、アドレスレジスタとデータレジスタの2つのレジスタをマッピングしています。アドレス・レジスタには、読み書きしたい特定のレジスタのインデックスが格納され、データ・レジスタにはそのレジスタのデータが格納されます。このことは、いくつかの例を見ることでより明確になるでしょう。

これらの部品を見る前に、VGAハードウェアの複雑さを強調したいと思います。VGAの詳細を網羅した本が何冊も出版されていますが、1つの記事ですべての詳細を網羅するのは非常に困難な作業です。ここでは、各コンポーネントの機能と、そのコンポーネントが使用するレジスタのセットについてのみ説明します。レジスタの数が多いため、ここではレジスタの説明はしません。残念ながら、ハードウェアレジスタの知識がないと、ビデオモードの設定方法を説明することができません。この問題を解決するために、すべてのレジスタのリストと、ビデオモードの設定の表を示します。

を最後に記載しています。これにより、読者は各レジスタの詳細やフォーマットを気にすることなく、VGAのビデオモードを設定することができる。ただし、VGAに興味のある方は[3]や[4]を参考にして、各レジスタの詳細を読むことをお勧めします。

要するに、レジスタの詳細はあまり気にしなくていい。ビデオモードリストを見れば、どのレジスタにどのような値を入れればよいかがわかります。

CRTコントローラー (CRTC)

Warning. **Improperly configuring the CRT Controller (CRTC) can potentially damage the video card or attached monitor.** Although rare, instances of CRT and LCD monitors have been known to burn out or explode due to improper or erroneous configuration.

CRTコントローラー (CRTC) は、ディスプレイ・モニターへのビデオ・データの出力を制御する役割を担っています。CRTCは、アドレスレジスタとデータレジスタによってアクセスされます。アドレスレジスタは3D4hに、データレジスタは3D5hにあります (Miscellaneous Output Register I/O Address selectフィールドが設定されている場合は、それぞれ3B4h、3B5h)。アドレス・レジスタを使用して、アクセスしたいCRTCレジスタを選択します。その後、データレジスタを使用して、そのCRTCレジスタから読み書きすることができます。

レジスターの数が多いため、各レジスターの完全な技術的レビューは行いません。CRTレジスタの詳細については、FreeVGAプロジェクトのページをご参照ください。ご要望があれば、将来的にはVGAに関するトピックを拡張するかもしれません。

CRTCレジスタは、ディスプレイの水平・垂直方向のリトレース期間とブランкиング期間のピクセルクロックのタイミングを制御することができます。これらのタイミングは、ディスプレイの出力 (ビデオ解像度) やリフレッシュレートの制御に役立ちます。その他のCRTCレジスタでは、ビデオメモリのスタートアドレスの変更 (プリセットロースキャンドラインとスタートアドレスレジスタ、ハードウェアカーソルの位置 (カーソルロケーションレジスタ) とアンダーライン (アンダーラインロケーションレジスタ)) 、CRTCモードコントロールレジスタでは、CRT自体といくつかのアドレスシングモードを制御することができます。

グラフィックスコントローラ

グラフィックスコントローラは、CPUとビデオメモリのインターフェースを管理する役割を担っています。アドレスレジスタは3CEhに、データレジスタは3CFhにマッピングされています。グラフィックスコントローラのレジスタにアクセスするには、アドレスレジスタにレジスタのインデックスを書き込み、データレジスタからリードまたはライトします。

以下に、標準的なレジスタの一覧を示します。各レジスタの詳細な説明については、FreeVGAプロジェクトをご参照ください。

シーケンサー

シーケンサーは、ビデオデータとRAMDACの間のインターフェースを管理します。アドレス・レジスタは3C4hに、データ・レジスタは3C5hにマッピングされています。シーケンサのレジスタにアクセスするには、アドレス・レジスタにレジスタのインデックスを書き込み、データ・レジスタからリードまたはライトします。

以下に、標準的なレジスタの一覧を示します。各レジスタの詳細な説明については、FreeVGAプロジェクトをご参照ください。

属性コントローラ

属性コントローラは、VGA用の21個のレジスタで構成されています。コントローラーには、I/Oポートスペースにマッピングされた2つのレジスターがあり、1つは3C0h、もう1つは3C1hにある。他のコントローラとは異なり、アトリビュートコントローラでは、3C0hのレジスタをデータ書き込みポートとアドレスポートの両方として使用する。3C1hのレジスタは、データリードレジスタです。アトリビュートコントローラーと通信するためには、ソフトウェアはまずポート3C0hにレジスタのアドレスを書き込み、続いてそのレジスタに書き込むデータを書き込まなければなりません。データの読み出しへは、ポート3C0hに書き込んだ後、3C1hから読み出すことで可能です。また、アトリビュートアドレスレジスタには特定のフォーマットがあります。

Attribute Address Register							
7	6	5	4	3	2	1	0
	PAS	Address					

アドレスとは、アクセスするためのレジスタ・インデックスのことです。以下のレジスター一覧表をご参照ください。

PAS (Palette Address Source) ホストまたはEGAディスプレイアダプタによるパレットデータへのアクセスを決定します。0であれば、ホストはパレットRAMにアクセスでき、アダプタはディスプレイメモリがパレットにアクセスすることを禁止します。1の場合、ディスプレイ・メモリはパレットRAMにアクセスでき、ホストはアクセスを禁止されます。

アトリビュートコントローラには、カラーインデックスと色を対応させる16個のパレットレジスタのセットが格納されている。このパレットの小ささはEGAの欠点である。VGAでは、これらのレジスタを引き続き使用するが、パレット情報を保存する代わりに、2番目のカラーインデックスレジスタのセットにアドレスを保存する。ここでは、その仕組みの詳細を説明しませんが、興味のある方は[4]394ページをご覧ください。

以下に、標準的なレジスタの一覧を示します。各レジスタの詳細な説明は、FreeVGAプロジェクトまたは[4]を参照してください。

汎用レジスター／外部レジスター

これらのレジスラリストは、まだまだ終わらないようですね。VGAの主要なコントローラのレジスラリストを見てきましたが、ビデオモードで使用される一般的な使用と雑多なデータのためのレジスタのセットはもっとあるので、見ておく必要があります。これらのレジスタは、一般的なレジスタまたは外部レジスターと呼ばれています。

以下に、標準的なレジスタの一覧を示します。各レジスタの詳細な説明は、FreeVGAプロジェクトまたは[4]を参照してください。

なお、カラーアダプタでVGA (EGAではない) を想定しています。モノクロのアダプタとEGAのアダプタでは、使用するポートが異なる場合があります。

カラーレジスター

これらのレジスターは、我々ソフトウェアが256色のパレットを操作・管理するためのものです。256色モードで作業する場合には重要なレジスター群であり、ビデオモードを設定した後に使用することもあるので、よく理解しておくとよいでしょう。また、これがこれから見る最後のレジスター群であることを知っておくとよいでしょう。

以下に、標準的なレジスターの一覧を示します。各レジスターの詳細な説明は、FreeVGAプロジェクトまたは[4]を参照してください。

標準ビデオモード

ビデオハードウェアを設定することで、異なる特性を持つさまざまな表示モードを定義することができます。しかし、ビデオハードウェアがサポートする多くのモードは、多くのモニターがサポートしていません。また、他の設定が望ましくない効果をもたらすこともあります。いくつかのVGAモードの設定は標準となっており、ほとんどのモニターでサポートされています。以下に、標準的なビデオモードとその構成の一覧を示します。標準的なビデオモードの番号は、モード0h、1h、2h、3h、4h、5h、7h、Dh、Eh、Fh、10h、11h、12h、13hです。その他の注目すべきモード、例えばモードXやモードQはよくサポートされていますが、標準モードではありません。

その他のリストは[3]または[4]を参照してください。以下の表は[4]の304～305ページから採用しました。この表は、各モードが変更するレジスタの異なるセットに分かれています。一番上の行はモード番号、左の列は特定のレジスタにアクセスするためのインデックスです。[4]によると、この値は標準的なビデオBIOSが使用する初期のモード状態を示しています。

ビデオモードを設定するには、ソフトウェアがビデオ出力を無効にし、変更するすべてのハードウェアレジスタに関連する値を書き込む必要があります。これにより、読み取り/書き込みモード、アドレスの開始、水平/垂直プランギング期間とタイミング、リフレッシュレート、解像度、グラフィックモードの種類など、すべてが設定されます。

正確さには万全を期していますが、以下の表には気づかなかつた誤りがあるかもしれません。この表は[4]で発表されたものと一致するように書かれていますが、書き換えの際に誤りがあるかもしれません。この点については、原著者に敬意を表します。

汎用レジスタ

Mode															
Index	0	1	2	3	4	5	6	7	D	E	F	10	11	12	13
0	63	63	63	63	63	63	63	A6	63	63	A2	A3	E3	E3	63
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70
3	4	4	5	5	4	4	5	FF	4	4	FF	4	4	4	4

Sequence registers

Mode															
Index	0	1	2	3	4	5	6	7	D	E	F	10	11	12	13
0	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	9	9	1	1	9	9	1	0	9	1	1	1	1	1	1
2	3	3	3	3	3	3	1	3	F	F	F	F	F	F	F
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	2	2	2	6	2	6	6	6	6	6	6	E

CRTC Registers

Mode															
Index	0	1	2	3	4	5	6	7	D	E	F	10	11	12	13
0	2D	2D	5F	5F	2D	2D	5F	FF	2D	5F	FF	5F	5F	5F	5F
1	27	27	4F	4F	27	27	4F	FF	27	4F	FF	4F	4F	4F	4F
2	28	28	50	50	28	28	50	FF	28	50	FF	50	50	50	50
3	90	90	82	82	90	90	82	FF	90	82	FF	82	82	82	82
4	2B	2B	55	55	2B	2B	54	FF	2B	54	FF	54	54	54	24
5	A0	A0	81	81	80	80	80	FF	80	80	FF	80	80	80	80
6	BF	FF	BF	BF	FF	BF	B	B	BF						
7	1F	FF	1F	1F	FF	1F	3E	3E	1F						
8	0	0	0	0	0	0	0	FF	0	0	FF	0	0	0	0
9	C7	C7	C7	C7	C1	C1	C1	FF	C0	C0	FF	40	40	40	41
A	6	6	6	6	0	0	0	FF	0	0	FF	0	0	0	0
B	7	7	7	7	0	0	0	FF	0	0	FF	0	0	0	0
C	0	0	0	0	0	0	0	FF	0	0	FF	0	0	0	0
D	0	0	0	0	0	0	0	FF	0	0	FF	0	0	0	0
E	0	0	0	0	0	0	0	FF	0	0	FF	0	0	0	0
F	31	31	59	59	31	31	59	FF	31	59	FF	59	59	59	31
10	9C	FF	9C	9C	FF	83	EA	EA	9C						
11	8E	FF	8E	8E	FF	85	8C	8C	8E						
12	8F	FF	8F	8F	FF	5D	DF	DF	8F						
13	14	14	28	28	14	14	28	FF	14	28	FF	28	28	28	28
14	1F	1F	1F	1F	0	0	0	FF	0	0	FF	F	0	0	40
15	96	96	96	96	96	96	96	FF	96	96	FF	63	E7	E7	96
16	B9	FF	B9	B9	FF	BA	4	4	B9						
17	A3	A3	A3	A3	A2	A2	C2	FF	E3	E3	FF	E3	C3	E3	A3
18	FF														

Graphics Controller

	Mode														
Index	0	1	2	3	4	5	6	7	D	E	F	10	11	12	13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	10	10	10	10	30	30	0	10	10	0	0	10	0	0	40
6	OE	OE	OE	OE	OF	OF	OD	OA	5	5	5	5	5	5	5
7	0	0	0	0	0	0	0	0	0	F	5	0	5	F	F
8	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF

Attribute Controller

	Mode														
Index	0	1	2	3	4	5	6	7	D	E	F	10	11	12	13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	13	13	17	8	1	1	8	1	3F	1	1
2	2	2	2	2	15	15	17	8	2	2	0	2	3F	2	2
3	3	3	3	3	17	17	17	8	3	3	0	3	3F	3	3
4	4	4	4	4	2	2	17	8	4	4	18	4	3F	4	4
5	5	5	5	5	4	4	17	8	5	5	18	5	3F	5	5
6	6	6	6	6	6	6	17	8	6	6	0	14	3F	14	6
7	7	7	7	7	7	7	17	8	7	7	0	7	3F	7	7
8	10	10	10	10	10	10	17	10	10	10	0	38	3F	38	8
9	11	11	11	11	11	11	17	18	11	11	8	39	3F	39	9
A	12	12	12	12	12	12	17	18	12	12	0	3A	3F	3A	0A
B	13	13	13	13	13	13	17	18	13	13	0	3B	3F	3B	0B
C	14	14	14	14	14	14	17	18	14	14	0	3C	3F	3C	0C
D	15	15	15	15	15	15	17	18	15	15	18	3D	3F	3D	0D
E	16	16	16	16	16	16	17	18	16	16	0	3E	3F	3E	0E
F	17	17	17	17	17	17	17	18	17	17	0	3F	3F	3F	0F
10	8	8	8	8	1	1	1	OE	1	1	OB	1	1	1	41
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	OF	OF	OF	OF	3	3	1	OF	OF	OF	5	OF	OF	OF	OF
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ビデオモードを設定するには、上の表にある、モードが変更するレジスタをすべて設定する必要があります。以下の例では、モードの設定を説明しています。

レジスタの値を書き込んだ後、ビデオモードの変更に成功しました。



VGAモード13hで動作し、ディスプレイメモリをクリアするデモ

- 3.** 以前、平面メモリモデルを見たとき、ハードウェアレジスタを見るまでは、ピクセルをプロットする例を完成させることができなかったことを思い出してください。ここではその例を紹介します。

4. SuperVGA Interface

SuperVGAとは、SuperVGA (SVGA) をはじめ、XGA、SXGA、SXGA+、UXGA、QXGA、QSXGAなど、2560x2048以上の解像度に対応した規格を含むディスプレイハードウェアのクラスを指す。SuperVGAは、メーカーがVGAを拡張する形で始まった規格で、現在のSuperVGAとは別の規格となっています。つまり、SuperVGA機器のすべてを網羅した規格は存在しないのだ。各SuperVGAカードはそれぞれ異なり、ハードウェアやファームウェアのインターフェースも異なります。そこで、問題が発生します。標準的な使い方がないのに、どうやってSuperVGAをサポートするのか？この問題は、ほとんどのメーカーがカードの技術仕様を公開していないことで、さらに深刻化しています。その代わりに、WindowsやLinuxなどの異なるOS用にブラックボックスのドライバーを書いています。

1. このため、SuperVGAデバイスに対応するためには、2つの選択肢しかありません。
 2. 対応したいディスプレイデバイスの種類ごとに、デバイスドライバを書きます。
 3. Use Vesa Bios Extensions (VBE)

1つ目の方法は、仕様書がないと作業ができないで難しいです。仕様書入手できないデバイスの場合、リバースエンジニアリングに頼るしかありません。今回のSuperVGA「カード」は、Bochsエミュレータで使用されているものを使用します。これを選んだのは、Bochsを使っている人なら誰でも使えるからです。ただし、ドライバのコードはBochs専用となります。

2つ目の方法は、VBE (Vesa Bios Extensions) を使用することです。VBEは、SuperVGAハードウェアのためのBIOS割り込み拡張の標準セットです。VBEは、非常にシンプルで、SuperVGAを扱うための単一の標準的なインターフェースを実現しています。ただし、リアルモードまたはv8086モードが必要で、BIOS割り込みを使用するため、すべてのマシンがVBEをサポートしているわけではありませんし、拡張機能でもありません。

3.1. Vesa Bios Extensions (VBE) Firmware Interface

VESA Bios Extensions (VBE) は、SuperVGAモードで動作するためのBIOS割り込みサービスを定義しています。VBEは、VESA Bios Extensions (VBE) Core Standardで定義されています（参考文献[5]）。（カーネルモードのソフトウェアは、リアルモードまたはv8086モードでBIOSサービスを呼び出すことができます。

VBEモード番号

VBEでは、ビデオBIOSと同様に、モード番号を定義しています。モード番号は次のような形式になっています。

VBE Mode Number							
15	14	13	12	11	10	9	Bits 0...8
DM	LFB	Reserved, set to 0			Mode		

Mode numberは、実際のモード番号です。ビット8がセットされていれば、VESA定義の標準モードです（詳細は後述します）。

LFBは、モードをLFB (Linear Frame Buffer) またはバンクスイッキングモードとして選択します。0の場合、バンクスイッキングを使用し、標準的なVGAフレームバッファを使用します。1の場合は、リニアフレームバッファを使用します。

DMは、モード設定時にディスプレイメモリをクリアするかどうかを選択します。0であれば、ディスプレイメモリをクリアします。1の場合は、ディスプレイメモリをクリアしません。

VESA定義のモードは、BIOSベンダーがサポートを推奨するビデオモード番号の標準セットです。次の表は、グラフィックモードの一覧です。

Mode	Resolution	Color depth	Mode	Resolution	Color

				depth
100h	640x480	256	113h	800x600
101h	640x480	256	114h	800x600
102h	800x600	16	115h	800x600
103h	800x600	256	116h	1024x768
10Dh	320x200	32K	117h	1024x768
10Eh	320x200	64K	118h	1024x768
10Fh	320x200	16.8M	119h	1280x1024
110h	640x480	32K	11Ah	1280x1024
111h	640x480	64K	11Bh	1280x1024
112h	640x480	16.8M		

3.2. 11Bhを超えるモード番号も定義できますが、標準ではありません。そのため、より高い解像度のモードをサポートすることも可能です。

VBEサービス

次に、VBE BIOSのサービスについて説明します。これらのサービスはすべて[5]で参照できますので、覚えておいてください。ここでは、ビデオモードの設定とディスプレイのメモリアクセスに必要な、最も重要な3つのサービスについてのみ説明します。その他の割り込みについては、仕様書（読みやすい仕様書の一つです）を参照してください。

INT 0x10 Function 4F00h - VBE Controller Informationの取得

入力します。

AX=4F00h

ES:DI=VbeInfoBlock構造体へのポインタ(以下の例を参照)

出力します。

AX=ステータス

Capabilities			
Bits 31...3	2	1	0
Reserved	D2	D1	D0

以下の例では、この割り込みを呼び出すことができます。

```

        return;

/* call BIOS */
in.eax.val = 0x4F00;
in.es = SEG((unsigned int) descr);
in.edi.val = OFFSET((unsigned int) descr);
io_services (0x10, &in, &out);
}

```

INT 0x10 Function 4F01h - Get VBE Mode Info**入力します。**

AX=4F01h

CX=モード番号(VBEのモード番号のフォーマットを思い出してください！)

ES:DI=ModeInfoBlock構造体へのポインタ(以下の例を参照)

出力します。

AX=ステータス

次の例では、割り込みの呼び出しを行っています。

Example. The following function uses the above interrupt for basic bank switching. Note the use of **SEG** and **OFFSET** to convert the 32 bit linear address of **out** to a 16 bit segment:offset far pointer.

```

void vbe_get_mode (int mode, modeInfoBlock* descr) {
    INTR in, out;

/* sanity check */
if (!descr)
    return;

/* call BIOS */
in.eax.val = 0x4F01;
in.ecx.val = mode;
in.es = SEG ((unsigned int)descr);
in.edi.val = OFFSET((unsigned int)descr);
io_services (0x10, &in, &out);
}

```

INT 0x10 Function 4F02h - Set VBE Mode

最後に紹介するのは、ディスプレイモードを設定するためのインターブトです。これは、VGAやVBEで定義されたSuperVGAモードや、標準ではない拡張モードの設定に使用できます。

入力します。

AX=4F02h

BX=モード番号 (VBEのモード形式を覚えておこう！)。

出力します。

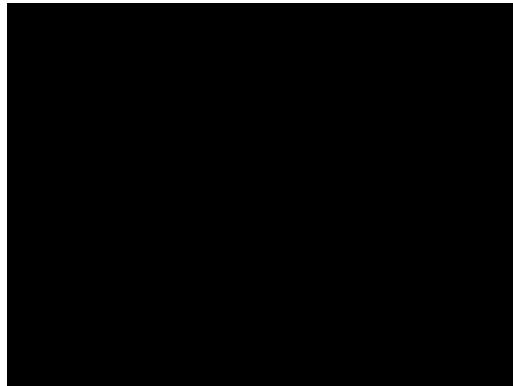
AX=ステータス

以下の例では、VBEモードを設定しています。

◆ Example. The following function uses the above interrupt to set a VBE mode.

```
void vbe_set_mode (int mode) {
    /* call BIOS */
    INTR in, out;
    in.eax.val = 0x4F02;
    in.ebx.val = mode;
    io_services (0x10, &in, &out);
}
```

vbe_set_modeを呼び出すと、VBEで定義された任意のモードを設定することができ、次のような結果が得られます。



SuperVGA Mode 118h (1024x768)で動作するデモ

VBEディスプレイメモリ

VBEは、SuperVGAで採用されている2つの標準的な表示メモリの方式に対応しています。リニアフレームバッファ（LFB）」と「バンクスイッチング」だ。VBEがサポートするすべてのビデオモードをこの2つで設定しようとしていることを思い出してください。これらは表示メモリへのアクセス方法の違いに過ぎません。この章では、両方のモードを扱う方法について説明します。

リニアフレームバッファ（LFB）モード

LFBモードでは、すべてのビデオメモリが物理アドレス空間にマッピングされ、通常は3GB～4GBの範囲の上位アドレスにマッピングされます（IA32アーキテクチャを想定）。ディスプレイメモリはC言語の配列のように直線的で、ポインタから読み書きするだけで、すべてのビデオメモリにアクセスできます。VBEでは、INT 0x10 Function 4F01hを呼び出して、現在のモードのmodeInfoBlock構造を取得することで、このポインタを得ることができます。その時に表示するポインタは、ちょうどmodeInfoBlock.physbaseになります。

LFBモードのディスプレイメモリは、マッピングされているディスプレイメモリの量が多いため、プロテクトモードまたはロングモードからしかアクセスできません。リアルモードでは、高解像度モードのディスプレイメモリにアクセスするには、バンクスイッチを使用するしかありません。

バンク切り替え

バンク切り替えを使用するモードでは、常に64KのディスプレイメモリがVGAメモリスペースの0xa0000の物理アドレスにマッピングされています。この64Kのディスプレイメモリのブロックをバンクと呼ぶ。ソフトウェアは一度に1つのバンクにしかアクセスできないため、ディスプレイメモリのすべてにアクセスすることはできない。すべてのディスプレイメモリにアクセスするためには、必要に応じてバンクを切り替える必要があります。

つまり、バンク0はディスプレイメモリの0～64Kバイト、バンク1はディスプレイメモリの64K～128Kバイトといった具合である。つまり、バンクを切り替えることで、ソフトウェアはこの64Kの"窓"を介して、どんなに高解像度のディスプレイメモリにもアクセスできるのだ。

バンク切り替えモードは、より多くの計算を必要とするため、一般的にはより遅くなります。つまり、画面上のピクセルの位置だけでなく、バンク内のオフセットや切り替えるバンクも計算しなければなりません。この計算は、ディスプレイメモリからの読み出し時にも同じです。しかし、バンク切り替えモードは、64Kのディスプレイメモリをマッピングするだけなので、リアルモードやv86モードでも使用することができます。

VBEでは、ソフトウェアはINT 0x10 Function 0x4F05 - Display Window Controlを呼び出してバンクを切り替えることができました。また、INT 0x10 Function 0x4F01 - Get VBE Mode Informationを呼び出し、VbeModeInfo.WinFuncPtrを呼び出すことで、（VBE仕様で推奨されているように）直接呼び出すこともできます。

INT 0x10 Function 4F05h - ディスプレイウィンドウコントロール

ディスプレイバンクの設定または取得を行います。VbeModeInfo.WinFuncPtrからも呼び出せます。

入力します。

AX=4F05h

BH=0 (メモリウィンドウの設定), 1 (メモリウィンドウの取

得) BL=0 (ウィンドウA), 1 (ウィンドウB)

DX=ウィンドウの粒度単位でのウィンドウ番号。BH=0 (セット機能) の時のみ使用

出力します。

AX=ステータス

5. DX=グラニュラリティ単位のウィンドウ番号。BH=1 (Get機能) の時のみ使用

6. Bochs VBE Interface

Bochsエミュレータは、ファームウェアを呼び出さずにVBEモードを直接設定する代替方法を提供します。サポートは限られていますが、高解像度のグラフィックスを扱うためのセミポータブルな方法（互換性のあるBochsエミュレータのみが必要）を提供してくれます。

```
B1Bochsvbewrite (VBE_DISPI_INDEX_ENABLE, VBE_DISPI_ENABLED | VBE_DISPI_LFB_ENABLED);
```

標準的なスーパーVGAのハードウェアインターフェースは存在しないことを思い出してください。つまり、この方法はBochsに特化したものであり、他の環境やプラットフォームでは動作しない可能性があります。読者の皆様は、Bochsをメインのエミュレータとしてお使いになると思いますので、最高の互換性を保つために、この方法を採用します。また、この方法は最もシンプルな方法でもあります。

しかし、互換性を高めるためには、VBEのサービスを直接利用するために仮想8086モードを使用することをお勧めします。

7. Demos

今回の記事で公開を予定しているデモは複数あります。VGA BIOSデモ、VBE SuperVGAデモ、VGAハードウェアデモの3つです。

8. Conclusion

この章では、VGA BIOS、VGAハードウェア、SuperVGA、VESA BIOS Extensionsの紹介など、多くのことを説明しました。これで、後の章で必要となる低解像度と高解像度のグラフィックモードの切り替えに必要な資料をすべて網羅しました。次の数章では、実際のグラフィックレンダリングとバイオペインに焦点を当て、SuperVGAについても触れていきます。次回は、基本的なブリミティブとグラフィックレンダリング、そしてトランシスフォームから始める予定です。

後のデモで使用する方法は、今回最後に紹介した「Bochs VBE Interface」です。これにより、読者の多くがBochsを使用すること前提に、後の記事で高解像度のLFBモードを使用することができます。つまり、どのような方法でディスプレイモードを設定しても、グラフィックスに焦点を当たったこれからの数回の記事に従うことができるのです。

次の機会まで。

～Mike () です。

OS開発シリーズ編集部

6. Resources

[1] 以下のリンクは、より直接的で正確な情報を提供するために参照されたものです。さらなる情報を得るために参照してください。

[2] http://en.wikipedia.org/wiki/Video_Graphics_Array

[3] <http://www.osdever.net/FreeVGA/vga/vga.htm>

[4] http://wiki.osdev.org/VGA_Hardware

[5] Programmer's Guide to the EGA, VGA, and Super VGA Cards (3rd Edition)

[6] VESA Bios Extensions (VBE) 2.0 Standard

[Home](#)



Operating Systems Development Series

オペレーティングシステム開発 - マルチブート

by Mike, 2010

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

はじめに

このチュートリアルでは、マルチブート規格と、マルチブート・コンプライアント・オペレーティング・システムの開発方法について説明します。このシリーズでは、マルチブートの構造について説明しますが、マルチブート・コンプライアント・システムを作成するには、それ以上のことが必要です。あなたのシステムがマルチブート・コンプライアントであれば、どのマルチブート・コンプライアント・ブートローダでもロード可能になります。すげいでしょ？つまり、どのマルチブート・コンプライアント・ブートローダでも、あなたのOSを起動できるということです。

マルチブート仕様

アブストラクト

マルチブート規格は、1995年に作成されたもので、Free Software Foundationが監修しています。マルチブートを可能にする標準的な方法を定義した書面による仕様を提供している。マルチブートとは、コンピュータシステムに複数のオペレーティングシステムやシステム環境をインストールして実行することができるものです。デュアルブートコンピュータシステムは、2つのオペレーティングシステムがインストールされているマルチブートの一例です。

マルチブートを可能にするのは、もうひとつの非公式なソフトウェアの仕掛けです。パーティショニングです。パーティショニングとは、1つの物理ディスク上に複数の論理ディスクが存在するような状態を作り出すことです。パーティショニングは、記憶媒体（通常はハードディスク）上のストレージを異なる用途に分けます。例えば、最初のパーティションはセクター0からセクターnまであり、NTFSフォーマットのWindows OSが入っています。他のパーティションには、データや他のオペレーティングシステムソフトウェアを含む、任意のファイルシステムを使用できます。

パーティション設定はソフトウェアのトリックなので、ブートローダがソフトウェアのパーティションテーブルを読み込んでこれらのパーティションを検出し、通常、起動するオペレーティングシステムが入っているパーティションのリストを表示できるかどうかにかかっています。これがブートメニューです。マルチブート仕様では、ブートローダがOSに制御を移すときのコンピュータの状態や、OSへのデータの受け渡し方法などを定義しています。マルチブート規格は、マルチブートに対応していないディスクにも使用することができます。つまり、マルチブートのコンプライアントブートローダがフロッピーディスクから起動できる場合、フロッピーディスクのOSをそこから起動させることができます。

オペレーティングシステムイメージ

一般的なブートローダは、さまざまな種類のOSイメージを起動するように設定できます。一般的には、これはカーネルや他のOS読み込みプログラムです。マルチブート仕様では、イメージのフォーマットについての詳細は規定されていません。このため、フラットバイナリ、ELF、あるいはPEファイルなど、好きなフォーマットを使用することができます。

ただし、このファイルには、マルチブートヘッダーという追加のヘッダーが必要です。このヘッダーは、イメージの最初の8kのどこかに、ワード（32ビット）単位で配置されていなければなりません。マルチブートコンプライアントブートローダは、このヘッダを見つけて情報を得ることができます。これにより、ブートローダはあなたのイメージをどのようにロードして実行するかを知ることができます。

ここでは、構造体のフォーマットを紹介します。

```
typedef struct _MULTIBOOT_INFO {  
    uint32_t magic;           //all required...  
    uint32_t flags;
```

```

    uint32_t checksum;
    uint32_t headerAddr; //all optional, set if bit 16 in flags is set...
    uint32_t loadAddr;
    uint32_t loadEndAddr;
    uint32_t bssEndAddr;
    uint32_t entryPoint;
    uint32_t modType; //all optional, set if bit 2 in flags is set...
    uint32_t width;
    uint32_t height;
    uint32_t depth;

} MULTIBOOT_INFO, *PMULTIBOOT_INFO;

```

パディングが追加されていないことを確認する必要があります。MSVCでは、上記の構造体宣言の周りに #pragma pack(push,1) と #pragma pack(pop,1) を追加することで、これを実現できます。

GRUBのようなマルチブート・コンプライアント・ブートローダでOSを起動するために必要な構造は、上記の通りです。メンバーを見てみましょう。

magic: 常に0x1BADB002でなければなりません。

- **flags:**

ビット 0

- 0: All boot modules and OS image must be aligned in page (4k) boundaries.

ビット 1

- 1: Boot loader must pass memory information to the operating system.

ビット 2

- 1: ブートローダーはビデオモードテーブルをOSに渡さなければならない。

ビット 16

1: マルチブートヘッダのオフセット12~28が有効です。(つまり、マルチブートヘッダのメンバー header_addr から entry_addr が有効です)。このビットがセットされていると、ブートローダはイメージフォーマットを解析してそこから値を取得する代わりに、これらの値を使用します。マルチブートコンプライアントブートローダは、ロードできるELFやPEなどのネイティブ実行ファイルフォーマットのサポートを提供できます。

チェックサム。これは、magicとflagに加えて、0の32ビット符号なしの合計でなければならない値でなければなりません。

headerAddr: マルチブートヘッダのアドレス

loadAddr: ロード先のベースアドレス

loadEndAddr: ロード終了アドレス。0の場合、ブートローダはOSイメージファイルの終わりと見なします

bssEndAddr: BSSセグメントの終わり。ブートローダはこのセグメントを無効にします。0の場合、BSSセグメントがないと判断されます

entryPoint: エントリポイント関数のアドレス。そうです、OSのエントリーポイントです。

modTypeです。

0: リニアグラフィックモード

1: EGA 標準テキストモード そ

の他はすべて予約済み

width: ディスプレイの幅をテキストの列またはピクセルで指定します。0の場合、ブートローダは設定を想定していません height: ディスプレイの高さをテキストの列またはピクセルで表します。0の場合、ブートローダは depth を設定しません。グラフィックスモードでのBPP(Bits Per Pixels)の数。0の場合、ブートローダは何も設定しません。

それがすべてです。ブートローダーは、2つの方法でOSをロードして実行することができます。1つは、実行可能なイメージフォーマット (ELFやPEなど) を読み込んで実行する方法、もう1つは、この構造体にある情報を利用する方法です。

ブートローダーは、イメージの中にあるこの構造を探します。そのため、この構造を記入して作成する必要があります。

マルチブートヘッダーの実装

マルチブート・ヘッダーをOSに組み込むには、さまざまな方法があります。ツールチェインごとに異なるソリューションがあります。そのうちのいくつかを見てみましょう。

Visual C++ 2005、2008

これは、私が最近見つけて他のフォーラムに投稿したトリックです。これは、Microsoft Visual C++が提供する拡張機能を利用して、カーネル内のヘッダを定義するものです。

まず、構造体を宣言し、余分なパディングがないことを確認します。

```
#pragma pack (push, 1)

<太后
*   Multiboot structure
*/
typedef struct _MULTIBOOT_INFO {

    uint32_t magic;
    uint32_t flags;
    uint32_t checksum;
    uint32_t headerAddr;
    uint32_t loadAddr;
    uint32_t loadEndAddr;
    uint32_t bssEndAddr;
    uint32_t entryPoint;
    uint32_t modType;
    uint32_t width;
    uint32_t height;
    uint32_t depth;

} MULTIBOOT_INFO, *PMULTIBOOT_INFO;

#pragma pack (pop, 1)
```

あとは、この構造体をどこかに定義するだけです。このヘッダーは、ワード（32ビット）バウンダリで、カーネルの最初の8K以内に定義しなければならないことを覚えていていますか？このトリックでは、構造体の適切な配置を保証するためにセクションアライメントを使用します。IDEのリンクオプションでセクションアライメントを設定すると、ワードアライメントになることが保証されます。あとは、新しいプログラムセクションを作り、その中に構造体を定義するだけです。すっきりしましたね。

今すぐにでもそうしましょう。

```
//! Bad example:
#pragma section(".text")
__declspec(allocate(".text"))
MULTIBOOT_INFO _MultibootInfo = {

    MULTIBOOT_HEADER_MAGIC,
    MULTIBOOT_HEADER_FLAGS,
    CHECKSUM,
    HEADER_ADDRESS,
    LOADBASIC,
    0, //load end address
    0, //bss end address
    KeStartup
};
```

これは動作しますが、問題があります。これは.textセクションに構造体を割り当てますが、どこに？マルチブートの仕様では、構造体をイメージの最初の8Kに配置することが要求されていますが、MSVCはまだ8K領域外に自由に配置することができますので、これは問題になります。

この問題を解決するには、セクションの命名規則を使う必要があります。MSVCで使われているセクションの命名規則は、name\$locというフォーマットに従っています。nameはセクションの名前で、locはセクション内のどこを表すかを示す英数字です。セクション\$aが1番目、セクション\$bが2番目というように、英数字の順に並んでいます。つまり、.text\$0を使えば、.textセグメントの先頭を表すことになります。しかし、もちろん、上記の.textを.text\$aに置き換えるだけではうまくいきません。）

これをコードセグメントとして作成し、.textセクションにマージすることができます。

```
//! Complete example
#pragma code_seg(".a$0")
__declspec(allocate(".a$0"))
MULTIBOOT_INFO _MultibootInfo = {

    MULTIBOOT_HEADER_MAGIC,
    MULTIBOOT_HEADER_FLAGS,
```

```

    CHECKSUM,
    HEADER_ADDRESS,
    LOADBASE,
    0, //load end address
    0, //bss end address
    KeStartup
};

#pragma comment(linker, "/merge:.text=.a")

```

それがすべてです。OABASEはカーネルのベースアドレス（例えば1MB）、HEADER_ADDRESSはマルチブートヘッダのアドレス（.textが常にオフセット0x400で始まるため、LOADBASE+0x400になります）、magicは0x1BADB002、flagsは0x00010003、チェックサムは-(MULTIBOOT_HEADER_MAGIC + MULTIBOOT_HEADER_FLAGS)です。

ここでは、その完全な例を紹介します。

```

#pragma pack(push, 1)

<**
 *  Multiboot structure
 */
typedef struct _MULTIBOOT_INFO {

    uint32_t magic; uint32_t
    flags; uint32_t
    checksum; uint32_t
    headerAddr; uint32_t
    loadAddr; uint32_t
    loadEndAddr; uint32_t
    bssEndAddr; uint32_t
    entryPoint; uint32_t
    modType; uint32_t
    width; uint32_t height;
    uint32_t depth.

}multiboot_info, *pmultiboot_info;

#pragma pack(pop,1)

<**
 *  Kernel entry
*/
void KeStartup ( PMULTIBOOT_INFO* loaderBlock ) {
    __halt () です。
}
// ! 読み込みアドレス
#define LOADBASE          0x100000
// ! ヘッダーオフセットは常にこのようになります
#define ALIGN             0x400
#define HEADER_ADDRESS    LOADBASE+ALIGN

#define MULTIBOOT_HEADER_MAGIC      0x1BADB002
#define MULTIBOOT_HEADER_FLAGS     0x00010003
#define STACK_SIZE                0x4000
#define CHECKSUM                 -(MULTIBOOT_HEADER_MAGIC + MULTIBOOT_HEADER_FLAGS)
#pragma code_seg(".a$0")
__declspec(allocate(".a$0"))

MULTIBOOT_INFO _MultibootInfo
= {

    multiboot_header_magic,

```

```
multiboot_header_flags,  
checksum, header_address,  
loadbase,  
0, //load end address 0,  
//bss end address  
KeStartup  
};
```

```
#pragma comment(linker, "/merge:.text=a")
```

このカーネルが1MBのベースアドレスを持ち、Visual C++でコンパイルして有効なPE実行ファイルを作成したと仮定すると、これはあらゆるマルチブート・コンプライアント・ブートローダでブート可能なはずです。

マシンの状態

ブートローダーが我々のOSを実行するとき、レジスタは以下の値になっている必要があります。

EAX - マジックナンバー。0x2BADB002でなければなりません。これはカーネルに、私たちのブートローダーがマルチブート規格であることを示します。

EBX - マルチブート情報構造体の物理アドレスが格納されています。

CS - オフセットが`0'、リミットが`0xFFFFFFFF'の32ビットの読み取り/実行コード・セグメントでなければなりません。正確な値は未定義です。

DS,ES,FS,GS,SS - 32ビットのリード/ライトデータセグメントでなければならず、オフセットは「0」、リミットは「0」です。

0xFFFFFFFF」となります。正確な値はすべて不定です。

A20ゲートが有効であること

CR0 - ビット31(PG)はクリア(ページング無効)、ビット0(PE)はセット(プロテクトモード有効)する必要があります。その他のビットは未定義

その他のレジスタはすべて未定義です。これらのほとんどは、既存のブートローダーすでに用意されています。追加しなければならないのは、EAXレジスタとEBXの2つだけです。私たちにとって最も重要なものはEBXに格納されています。これは、マルチブート情報構造体の物理アドレスを格納します。それでは見てみましょう。

マルチブート情報構造

さて、ブートローダによってOSが起動したところで、次は何をするのでしょうか？マルチブート・コンプライアント・ブートローダーは、OSに情報を提供する情報構造体も作成します。これらは、EBXレジスタの構造体へのポインタによって渡されます。

この構造体は、マルチブート仕様の中でも最も重要な構造体の一つです。この構造体の情報は、EBXレジスタからカーネルに渡されます。これにより、ブートローダがカーネルに情報を渡すための標準的な方法となります。

これはかなり大きな構造ですが、悪くはありません。これらのメンバーのすべてが必要なわけではありません。仕様では、オペレーティングシステムは、構造体のどのメンバーが存在し、どのメンバーが存在しないかを決定するために、flagsメンバーを使用しなければならないとされています。

ここでは、構造全体のフォーマットを示します。マルチブートのヘッダー構造と同様に、パディングがないことを確認することをお勧めします。

```
typedef struct _MULTIBOOT_INFO {  
    uint32_t flags; uint32_t          //required  
    memLower; uint32_t               //if bit 0 in flags are set  
    memUpper; uint32_t               //if bit 0 in flags are set  
    bootDevice; uint32_t              //if bit 1 in flags are set  
    commandLine; uint32_t             //if bit 2 in flags are set  
    moduleCount; uint32_t              //if bit 3 in flags are set  
    moduleAddress; uint32_t             //if bit 3 in flags are set  
    syms[4]; uint32_t                //if bits 4 or 5 in flags are set  
    memMapLength; uint32_t             //if bit 6 in flags is set  
    memMapAddress; uint32_t             //if bit 6 in flags is set  
    drivesLength; uint32_t              //if bit 7 in flags is set  
    drivesAddress; uint32_t             //if bit 7 in flags is set  
    configTable; uint32_t              //if bit 8 in flags is set  
    apmTable;                      //if bit 9 in flags is set  
    uint32_t vbeControlInfo;           //if bit 10 in flags is set  
    uint32_t vbeModeInfo;              //if bit 11 in flags is set  
    uint32_t vbeMode; uint32_t          // all vbe_* set if bit 12 in flags are set  
    vbeInterfaceSeg; uint32_t  
    vbeInterfaceOff;  
    uint32_t vbeInterfaceLength;  
} MULTIBOOT_INFO, *PMULTIBOOT_INFO;
```

この構造は、見た目ほど複雑ではありません。flagsメンバの対応するビットがセットされていれば、上に示したメンバが有効であることを意味します。このため、技術的には、flagsが唯一の必須メンバーであり、他のメンバーはすべてオプションです。

ここにいるメンバーを見てみましょう。

memLow, memUpper: 低い方のメモリと高い方のメモリの量（単位：KB）。下位メモリは0から、上位メモリは1MBから。

bootDevice。ブートデバイス（後述）

commandLine: カーネルのコマンドラインを含むC言語の文字列へのポインタ

moduleCount。ブートローダーでロードされた追加ブートモジュールの数

moduleAddress: 最初のモジュール構造体のアドレス（後述）

syms: シンボルテーブルの位置。以下を参照 memMapLength: システム

メモリマップのエントリ数 memMapAddress: メモリマップのアドレス

drivesLength, drivesAddress: 以下参照

configTable: BIOS ROMコンフィギュレーションテーブルのアドレス（GET CONFIGURATION BIOS INTコールから返される）

apmTable: アドバンストパワーマネージメント（APM）テーブルのアドレス

vbeControlInfo, VbeModeInfo: Video Bios Extensions（VBE）構造体のアドレスです。

vbeMode: VBEモード

vbeInterfaceSeg, vbeInterfaceOff, vbeInterfaceLength: VBE 2.0のプロジェクトモードのインターフェースにアクセスするために使用します。

本章ではVBEやAPMについては触れていませんので、ここでは割愛します。メモリーマップについては、第17章でシステムメモリーマップのフォーマットを含めて説明しました。

configTableのROM構成は、BIOS INT 0x15 Function 0xC0から得られるテーブルです。

それがすべてです。この構造は悪くないですね :) bootDevice、moduleAddress、syms、drivesLength、drivesAddressなど、まだ見ていないメンバーがいくつかあります。これらを詳しく見ていきましょう。

ブートデバイス

bootDeviceメンバは、以下のフォーマットに従います。1ワード目：BIOSドライブ番号

2語目、3語目、4語目 パーティション

BIOSのドライブ番号は、BIOSのINT 0x13サービスでドライブを表すために使用される番号です。その他の単語はパーティションを表します。単語2,3,4はパーティション1,2,3を表します。パーティション1はトップレベルのパーティションで、パーティション2はその中のサブパーティションです。未使用のパーティションは0xFFと表示されます。

moduleAddress

これは、最初のモジュール構造体へのポインタです。モジュール構造体のエントリーは、以下のフォーマットに従います。

```
typedef struct _MODULE_ENTRY {
    uint32_t moduleStart;
    uint32_t moduleEnd;
    char     string[8];
} MODULE_ENTRY, *PMODULE_ENTRY;
```

moduleStartとmoduleEndは、ロードされたモジュールの開始と終了のアドレスを含みます。string "はそのモジュールを表し、通常はコマンドライン名やパス名、または何もない場合は "0 "となります。

drivesLength, drivesAddress

drivesLengthメンバには、すべてのドライブ構造体のサイズが含まれています。 drivesAddressには、最初のドライブ構造体へのポインタが含まれています。 ドライブ構造体のエントリは、以下の形式を持ちます。

```
typedef struct _DRIVE_ENTRY {
    uint32_t size;           //size of structure
```

```

    uint8_t driveNumber;
    uint8_t driveMode;
    uint16_t driveCylinders;
    uint8_t driveHeads;
    uint8_t driveSectors;
    uint8_t ports [0];           //can be any number of elements
}DRIVE_ENTRY, *PDRIVE_ENTRY;

```

それでは、各メンバーを詳しく見ていきましょう。

driveNumber: BIOSで使用される番号

- **driveMode:**

0: CHS

1: LBA

ドライブシリンダー、ドライブヘッド、ドライブセクタ。ドライブジオメトリ

ports: ドライブへのアクセスにBIOSが使用するI/Oポート番号のリストを含み、0で終了します。

これがこの構造のすべてです。あと一人だけ、あの奇妙なsymsのメンバーを取り上げます...見てみましょう

syms

本章の構造体では、symsメンバはuint32_t syms[4]と宣言されていますが、これは完全には正しくありません。 実際にはいくつかのメンバーがあり、それらのメンバーに続くバイトを占有しています。

```

syms[0] = uint32_t sym_num syms[1]
= uint32_t sym_size syms[2] =
uint32_t sym_addr syms[3] =
uint32_t sym_shndx

```

仕様書には、OSイメージのフォーマットは何でもいい（ELF、PE、フラットバイナリなど）と書かれていますが、これはELFフォーマットに限った話です。技術的には、（カーネルのような）システムイメージは、それ自体を解析してシンボル情報を得ることができます。sym_numはELFセクションヘッダのシンボルエントリの数、sizeは各エントリのサイズ、addrはELFバイナリのシンボルテーブルのアドレスです。

結論

それがマルチブート規格のすべてです。技術的には、マルチブート・コンプライアント・ブートローダでシステムを起動するために、マルチブート・ヘッダーを適切に定義することだけが必要です。しかし、マルチブート情報構造を使って、通常はブート時に取得する情報を取得することができます。

シリーズでマルチブートをサポートしたい場合は、カーネルが仮想アドレスではなく物理アドレスでロードされるようにする必要があります。これは、マルチブート・コンプライアント・ブートローダがあなたのカーネルに制御を移すとき、ページングが無効になるからです。典型的なアドレスは1MBで、これはGRUBのような多くのブートローダでロードできます。もちろん、ページングを有効にして後から使用することもできます：）

次の機会まで。

~マイク

eptune Operating System Software

BrokenThorn Entertainment社。現在、「DoE」と「NSuite」を開発中。

ご質問やご意見がございましたらお気軽にお問い合わせください。





オペレーティングシステム開発シリーズ

IA32機械語

by Mike, 2011

はじめに

この章では、IA32機械語プログラミングについて説明します。ここで説明する内容は情報提供を目的としたものであり、基本的なオペレーティングシステムや実行ソフトウェアの開発には必要ありません。IA32（およびIA64）の命令フォーマットを理解することは、不適切にアセンブルされた命令のデバッグや、v8086モードをサポートするために必要なv86モニター、命令のエミュレーション（特定のFPU命令のエミュレーションや、アセンブラー、エミュレータ、仮想マシンなどの開発時に必要）、デバッガやコンバイラなどの特定のシステムソフトウェアの開発に役立ちます。

この章では、新章の執筆に使用する新しいエディターをテストするためのもので、フォーマットの改善やスペルミスの解消に役立つはずです。このテストが成功すれば、新章と旧章のすべてが新しいフォーマットに合わせて更新されます。もしエラーがあった場合は、ご意見をお寄せください。

機械語の概要

機械語は、マシンコード、ネイティブコード、バイトコードなどとも呼ばれ、中央処理装置（CPU）で実行可能な生の命令とデータのセットです。機械語は、CPUが特定のバイトシーケンスを、タスクを実行するための「命令」として解釈することを可能にする。これらのタスクは、少量のデータをコピーしたり、演算したりといった非常に小さなものです。CPUの命令を表すバイト列を構築する行為は、コーディングと呼ばれる。コーディングの定義は、プログラミング言語の進化に伴って変化してきた。当初は、命令のバイト列を実際にコーディングすることを指していたが、現在では、第2世代、第3世代、第4世代のプログラミング言語によるさまざまな形態のプログラミングを指す。コンピュータプログラムはソフトウェアとも呼ばれ、ワープロやHalo®のプレイなど、複雑な作業を行うための機械コードとデータの集合体である。一般的なメディアでは、機械語は「1と0の連続」と解釈されることが多い。これは、ある程度は正確な表現です。

デジタル・ロジック

デジタルロジックとは、論理ゲートを利用して電子機器に判断をさせる電子工学の分野です。論理ゲートの例としては、ANDゲート、ORゲート、NORゲート、NANDゲート、NOTゲート、XORゲートなどがある。これらのゲートは、2進法の演算を反映している。ANDゲートは2値のANDを、XORゲートは2値のXORを、というように。これらのゲートに意味を持たせるためには、何が真で何が偽なのかを理解するための規格を採用する必要があった。例えば、ANDゲートは、2つの入力と1つの出力がある場合にのみ意味を持つ。2つの入力は2つの項目で、どちらかが偽であれば出力は偽、そうでなければ出力は真となります。電流の少ない線を偽、電流の多い線を真と定義するのが定番だ。これが2進法とデジタルロジックの関係である。2進法では、0は偽、1は真と表記されることが多い。機械語が2進数で表現されることが多いのは、CPUの命令解読の仕組みや、RAMに格納して命令を取得する仕組みと密接に関係しているからだ。

プログラムの読み込み

プログラムは、オペレーティングシステム、エグゼキュティブ、またはファームウェアによってメモリにロードされます。IA32およびIA64ファミリーのCPUは、ROM (Read Only Memory) やRAM (Random Access Memory) からプログラムを実行することができます。これは、ファームウェアとプログラムイメージが共有するシステムバスと物理アドレス空間 (PAS) によって実現されている。ファームウェアとプログラムイメージは、どちらもCPUコアによって直接実行されるため、同じ機械語のバイトコードを使用します。機械語は、ファームウェアが使用するマイクロコード（第7章参照）とは異なるが、実際のファームウェアは機械語であることに変わりはない。

アセンブリ言語

アセンブリ言語は、第二世代のプログラミング言語です。アセンブリ言語では、プログラマーがソフトウェアを開発する際に役立つニーモニックを使用して、明確に定義された言語でプログラムを書くことができます。例えば、「MOV」は、さまざまなアーキテクチャに対応した多くのアセンブリ言語で共通のニーモニックです。MOVは、データをソースからデスティネーションにコピーする命令を表します。また、データ移動命令の一例もあります。

ニーモニックは、命令と命令形式に記号的な名前を与え、各アセンブリ言語命令を单一の（場合によっては複数の可能性のある）機械語バイトシーケンスに変換できるようにした。アセンブリ言語の命令を機械語に変換するプログラムは、アセンブラーと呼ばれる。アセンブラーは誤ってコンバイラと呼ばれることもある。

機械命令の概要

機械命令は、CPUの特定のタスクを実行する1バイトのシーケンスです。マシンインストラクションのセットは、以前はマシンランゲージとして定義されていた。機械語命令は、CPUメーカーが実装したすべてのCPU命令を記録した命令セットで定義されています。命令セットには、アセンブラー開発者が使用するための、示唆に富むアセンブリ言語のニーモニックも含まれているのが一般的です。命令セットはCPUの仕様書に記載されています。

CPUメーカーは、特定のCPUがサポートする機械命令を実装し、CPUが各命令をどのように解釈するかを決定します。これにより、CPUはメーカーが意図した機械語命令を「実行」することができます。しかし、CPUのハードウェアやファームウェアにバグがあると、有効な命令ではない命令をCPUが「実行」してしまうことがあります。これが文書化されていない命令です。アセンブラーによっては、よく知られている文書化されていない命令にニーモニックを定義している場合があります。文書化されていない命令の中には、後に実際の命令として文書化されたものもあります。IA32のLOADALL命令のように、メーカーのテスト用として文書化されていない命令もあります（このバグはその後修正されました）。また、システムを停止させたり、CPUにダメージを与えるような悪い影響を与える命令もあります（これらはHCF (Halt and Catch Fire) 命令として知られています）。

すべてのCPUアーキテクチャに対応した命令セットがあります。ソフトウェア業界の進化に伴い、命令セットにはある種の傾向が見られるようになりました。これらの傾向を理解することは、IA32の機械語を理解するのに役立ちます。

CISCとRISC

命令セットは、一般的に2つのカテゴリーに分類されます。CISC(Complex Instruction Set Computing)とRISC(Reduced Instruction Set Computing)です。RISCの例としては、PPCやARMのアーキテクチャがあります。CISCの例としては、IA32アーキテクチャがあります。RISCアーキテクチャは、CISCよりも単純な命令セット形式を採用しています。RISCアーキテクチャは通常、各命令に標準的なエンコーディングフォーマットを使用し、各命令が同じバイト数になるようにします。CISCアーキテクチャもまた、標準的なエンコーディングフォーマットに従いますが、可変長の命令が可能です。

操作コード

オペコード (OPCODE) とは、CPUが命令の種類を判断するために利用する1バイトの識別子です。例えば、MOV命令には、タイプ (MOV) などの命令に関する情報をCPUに知らせるオペコード識別子があります。多くの命令セットでは、1つの命令を別の命令と区別するためにオペコードを使用します。一部の命令は、マルチバイトオペコードや拡張オペコードを持つことがあります。これにより、命令セットの柔軟性が高まります。

アドレッシングモード

アドレスモードとは、CPUがアドレスを参照するための方法を定義するものです。アドレスは、アーキテクチャによって、仮想的なものと物理的なものがあります。データ移動命令などの命令は、データを取得するためにアドレスを参照する方法をCPUに伝える必要があります。例えば、多くのCPUはダイレクトアドレッシングモードをサポートしており、特定のアドレスのデータを参照（読み書き）するように命令がCPUに指示することができます。例えば、IA32のアセンブリ言語では

```
mov eax, dword [0xa0000].
```

この命令は、CPUに直接アドレッシングモードを使用して、現在のアドレス空間のアドレス0xa0000から読み取るように指示します。もうひとつの一般的なアドレッシングモードは間接アドレッシングで、ポインタを使ってデータを参照するようにCPUに指示することができます。例えば、IA32のアセンブリ言語では

```
mov eax, [ebp].
```

これにより、CPUはEBPレジスタに格納されているアドレスからEAXレジスタにドワードを読み込むことになります。この他にも、アーキテクチャによって様々なアドレッシングモードが存在します。

IA32 and IA64 Instruction encoding

さて、ここからはIA32とIA64の機械語命令のエンコーディングを見ていきましょう。ここでは、スペースを節約するために、IA32とIA64の命令セットという意味でIA64を使用します。IA32はIA64のサブセットであり、IA32の命令セットの大部分を共有しています。IA64の命令セットは、CISCエンコーディングを実装しています。これは、各命令が特定のエンコーディング構造に従っており、長さが可変であることを意味する。IA32とIA64の命令は、1バイトから12バイトまでの大きさがあります。

登録コード

CPUは内部のレジスタを数値で識別している。多くのレジスタは同じコードを共有していますが、CPUは使用する命令や現在の動作モード（リアルモード、プロテクトモード、ロングモード）に応じて使用するレジスタを決定します。使用するレジスタを決定する際には、オペランドサイズオーバーライドプレフィックスも使用されます。このプレフィックスについては後述します。

レジスタコードは、命令がどのレジスタを操作するのかをCPUに知らせるために、命令のエンコーディングで使用されます。レジスタには以下のコードが使われています。

REX.r =	0	1	2	3	4	5	6	7
<i>Code</i>								
No REX	AL	CL	DL	BL	AH	CH	DH	BH
REX	AL	CL	DL	BL	SPL	BPL	SIL	DIL
REG16	AX	CX	DX	BX	SP	BP	S1	D1
REG32	EAX	ECX	EDX	EBX	ESP	EBP	ES1	ED1
REG64	RAX	RCX	RDX	RBX	RSP	RBP	RS1	RD1
MM	MM0	MM1	MM2	MM3	MM4	MM5	MM6	MM/
XMM	XMM0	XMM1	XMM2	XMM3	XMM4	XMM5	XMM6	XMM7
YMM	YMM0	YMM1	YMM2	YMM3	YMM4	YMM5	YMM6	YMM/
SSEG	ES	CS	SS	DS	ES	GS		
CR0	CR1	CR2	CR3	CR4	CR5	CR6	CR/	
DR0	DR1	DR2	DR3	DR4	DR5	DR6	DR/	

例えば、mov bx, 0x5という命令では、BXのレジスタコードとして3を格納します。mov ss, axという命令では、SSのレジスタコードとして2を、AXのレジスタコードとして0を格納する必要があります。命令によって使用するレジスタの種類が異なるため、同じコードの複数のレジスタを選択しなければならないという矛盾が生じることはありません。例えば、mov REG16, IMM16という命令は、常に16ビットの汎用レジスタをオペランドとして使用します。また、movups xmm, xmm/m128という命令は、常にXMMレジスタのみを使用します。

ロングモードでは、このリストにさらにレジスタが追加されます。上記のレジスタをサポートするために、ロングモードでは、同じレジスタコードを使って他のレジスタを選択する命令を可能にする特別なフラグが設定されている。これが後に説明するREXプレフィックスバイトのREX.rフィールドである。このビットがセットされていると、レジスタテーブルは次のようになります。

REX.r=1	0	1	2	3	4	5	6	7
<i>Code</i>								
No REX	R8B	R9B	R10B	R11B	R12B	R13B	R14B	R15B
REX	R8W	R9W	R10W	R11W	R12W	R13W	R14W	R15W
REG16	R8D	R9D	R10D	R11D	R12D	R13D	R14D	R15D
REG32	R8	R9	R10	R11	R12	R13	R14	R15
MM	MM0	MM1	MM2	MM3	MM4	MM5	MM6	MM/
XMM	XMM8	XMM9	XMM10	XMM11	XMM12	XMM13	XMM14	XMM15
YMM	YMM8	YMM9	YMM10	YMM11	YMM12	YMM13	YMM14	YMM15

SSEG	ES CR8 DR8	CS CR9 DR9	SS CR10 DR10	DS CR11 DR11	FS CR12 DR12	GS CR13 DR13	CR14 DR14	CR15 DR15
------	------------------	------------------	--------------------	--------------------	--------------------	--------------------	--------------	--------------

命令のエンコーディング

IA64の命令は、CPUの8085に由来する明確な構造を持っています。各命令は次のような形式になっています。

Prefix bytes (0-4)	REX prefix (1)	Operation (0-3)	Mod R/M (1)	SIB (1)	Displacement (0-4)	Immediate (0-4)
--------------------	----------------	-----------------	-------------	---------	--------------------	-----------------

コンパクトするために、()内の数字はコンポーネントのバイト数です。0の数字は、そのバイトがオプションであることを示します。例えば、1つの命令の中で、プレフィックス・バイトは0から4バイトまであります。つまり、1つの命令には、0、1、2、3、4のプレフィックス・バイトが存在することになります。REXプレフィックスは、IA64およびlongモードでのみ有効です。唯一の必須フィールドはオペレーションコードです。他のフィールドはすべてオプションで、命令がそれらを必要とするかどうかに依存します。例えば、INT(割込み)命令では、操作コードと即値バイトが必要ですが、MOV命令では上記のフィールドをすべて使用することができます。

例として、INT命令をもう少し詳しく見てみましょう。INT命令には形があります。

INT imm8

ここで、imm8は8ビットの即値、INTは演算コード0xCDのニーモニックです。命令エンコードのフォーマットを知っていると、INT5命令を次のようにエンコードすることができます。

0xCD 0x05

1バイト目の0xCDはオペレーションコードで、茶色で表示されています。プレフィックスバイトはオプションであり、INT5では必要ないため、必要ありません。Mod R/MとSIBのバイトも必要ありません。ディスプレイスメントはメモリアドレッシングモードでのみ使用されるので、他に必要なフィールドは即値フィールドだけです。即時フィールドは、0~4バイトのフィールドです。INT imm8という命令形式があるので、1バイトのフィールドとして使うことができます。この例の目的は、あるフィールドはオプションであり、必要とされないことを示すことです。また、これらのフィールドの順序は決して変わりません。例えば、上の例では、不要なフィールドを省略することにしましたが、演算コードフィールドが即値フィールドの前にあるという、フィールドの順序はそのままです。次のセクションでは、これらの各フィールドについて詳しく説明します。

プレフィックス・フィールド

プレフィックス・バイトは、命令がCPUに対してより多くの情報を与えることを可能にします。例えば、CPUにバスをロックさせたり、データ移動命令で別のセグメントレジスタを使用させたりすることが可能になります。これらのプレフィックスの多くは、アセンブリ言語のニーモニックを持っています。プレフィックス・バイトは4つのクラスに分類されています。1つの命令で使用できるプレフィックス・バイトは、4つのクラスのそれぞれから最大1つです。

クラス1プレフィック

クス 0xF0 LOCK ブ

レフィックス

0xF2 REPNE, REPNZ 接頭辞 0xF3

REP, REPZ, REPE 接頭辞

クラス2プレフィックス

0x2E CS セグメントオーバーライド
0x36 SS セグメントオーバーライド
0x3E DS セグメントオーバーライド
0x26 ES セグメントオーバーライド
0x64 FS セグメントオーバーライド
0x65 GS セグメントオーバーライド

クラス3のプレフィックス

0x66 オペランドサイズオーバーライド

クラス4のプレフィックス

0x67 アドレスサイズオーバーライド

ここでは、読者がIA32アセンブリ言語を知っていることを前提としているので、これらのプレフィックスについての詳細な説明は省略します。1つの機械語命令には、4つのクラスのうち、1つのプレフィックス・バイトしか設定できません。4つのクラスがあるため、1つの命令は0~4個のプレフィックス・バイトを持つことができます。1つの命令が1つのクラスから2つ以上のプレフィックス・バイトを使おうすると、CPUは無効な命令の例外を発生させます。

LOCKプレフィックス

LOCKをサポートしていない命令にLOCKプレフィックスを使用した場合、CPUは無効な命令の例外を発生させます。アセンブリの中には、無効な命令に

LOCKを使用することをプログラマに警告せずに許すものがあります。そのため、ここでは、有効な命令のリストを紹介します。

LOCKプレフィックスは以下の命令でのみ使用可能です。ADC, ADD, AND, BTC, BTR, BTS, CMPXCHG, CMPXCHG8B, CMPXCHG16B, DEC, INC, NEG, NOT, OR, SBB, SUB, XADD, XCHG, XOR。

オペランドサイズオーバーライド

オペランドサイズオーバーライドは、CPUが16ビットと32ビットのオペランドを選択できるようにするものです。アセンブラーは通常、bits16やuse32などの指示を使って間接的にプログラマが特定のオペランドサイズを選択できるようにしています。IA32とIA64の命令セットでは、レガシーの16ビットとネイティブの32ビットの2種類のオペランドサイズが用意されています。ネイティブサイズは、プロセッサの現在の動作モードに依存します。

Operation mode	CS.d	REX.w	Native	Operand override
Real mode			16 bit	16 bit
V8086 mode			16 bit	16 bit
Protected mode	0		16 bit	32 bit
Protected mode	1		32 bit	16 bit
Long mode	0		32 bit	16 bit

例として、ADD AX/EAX, IMM16/IMM32命令を見てみましょう。この命令のオペレーションコードは0x05です。プロテクトモードのコードでは、CPUはデフォルトでこの命令をADD EAX, IMM32命令として解釈します。しかし、オペランドオーバーライドプレフィックスを使用することで、デフォルトの動作をオーバーライドし、16ビットの即値をコピーすることができます。アセンブリ言語では次のようにします。

```
add eax, 5 ; MOV EAX, IMM32
add ax, 5 ; MOV AX, IMM16
```

最初の指示では、組み立てを行います。

0 x 05 0 x 00 0 x 00

2回目の指示では、組み立てを行います。

0 x 66 0 x 05 0 x 00

この2つの命令の違いは、次の点だけです。(1) 最初の命令は32ビットの即値を使用し、2番目の命令は16ビットの即値を使用していること（これらは赤で表示）、(2) 2番目の命令はオペランドオーバーライドプレフィックスを使用していること（これは黒で表示）。これは、16ビットのオペランド形式を使用するようにCPUに指示するものです。念のため、茶色の値はオペレーションコードです。

アドレスサイズオーバーライド

アドレスサイズオーバーライドのプレフィックスバイトは、オペランドオーバーライドのプレフィックスバイトとよく似ています。アセンブリでは、byte ptrやdword ptrなどのキーワードを使って、プログラマがアドレスサイズを選択できるようになっています。この機能はオペランドオーバーライドプレフィックスと非常に似ているため、目的は同じですがアドレスモードに適用されるため、説明を省略します。

Operation mode	CS.d	REX.w	Native	Address override
Real mode			16 bit	16 bit
V8086 mode			16 bit	16 bit
Protected mode	0		16 bit	32 bit
Protected mode	1		32 bit	16 bit
Long mode	0		64 bit	32 bit
Long mode	1		64 bit	32 bit

例えば、mov eax, [0xa000]という命令をプロテクトモードでアセンブルした場合、アドレスサイズのオーバーライドは必要ありません。アセンブリは 0xa000 を 32 ビットの変位として扱います。しかし、mov ax, word [0xa000]を使用した場合、アセンブリは16ビットのアドレス形式を選択するために、命令にアドレスサイズオーバーライドのプレフィックスを追加します。

REXプレフィックス

REXプレフィックスは、64ビット特有の機能を有効にします。以下のような形式になっています。

	0		1	0		0	W R X B	0	
+	-	-	-	-	-	-	-	-	-
7									0

REX.w Operand size. 0: Default, 1: 64 bit
 REX.r ModRM.reg extension
 REX.x SIB.index extension
 REX.b ModRM.rm extension

プレフィックスオーダー

他のプレフィックスバイトと組み合わせて使用する場合、プレフィックスバイトの順序は重要ではありません。たとえば、マシンコードで0xF3 0x2Eを使用すると、REPとCSのオーバーライドを選択できます。また、0x2E 0xF3を使っても同じことができます。

操作コード欄

オペレーションコードフィールドは1~3バイトの長さになります。すべてのオペレーションコードはユニークで、使用する命令とそのオペランドを識別します。例えば、オペレーションコード0は、ADD REG/MEM8, REG8命令を示します。オペレーションコード1は、ADD REG/MEM16/MEM32, REG16/REG32命令を識別します。IA32およびIA64のCPUマニュアルに、各命令とそのオペレーションコードの概要が記載されています。

プライマリ Opcode

プライマリオペコードは、すべての命令で必要とされる1バイトです。命令を識別するためのオペレーションコードフィールドのベースとなるものです。一次オペコードは、命令によって次のような形式があります。

						d s w
+	-	-	-	-	-	-
7						0

					tttn	
7						0
					register ID	



PO.w Operand size
 PO.s Sign extend
 PO.d Direction
 PO.ttt Used on some FPU instructions
 PO.mf Memory format

セカンダリOPCコード

プライマリオペコードバイトが0xf0の場合、セカンダリオペコードバイトと呼ばれる別のバイトが続きます。セカンダリーオペコードは、異なる命令を識別するもので、上記と同じ機能を持っています。これらは2バイトのオペコードとして扱われます。

OPCode拡張

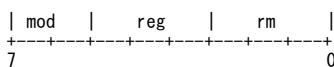
ある種の命令ファミリーは、同じオペコードを持ちながら、オペコードエクステンションと呼ばれる特別なフィールドによってのみ異なる。これは、Mod R/M.regフィールドに格納されている3ビットの拡張子です。Mod R/Mバイトについては、次のセクションで詳しく説明します。

マルチバイトOPCコード

プライマリオペコードフィールドの長さは1~3バイトです。ほとんどの命令はプライマリ・オペコード・フィールドの1バイトしか使用しませんが、中には3バイトすべてを使用するものもあります。これらの命令はすべて、セカンダリ・オペコード・バイト (0xf0) も使用します。

Mod R/Mフィールド

Mod R/M (Register/Memory) フィールドは、メモリまたはレジスタのオペランドを必要とする命令で使用されます。Mod R/Mフィールドのフォーマットは以下の通りです。



Mod R/Mフィールドは、CPUがリアルモード、プロテクトモード、ロングモードのいずれで動作しているかによって若干異なります。

Real mode

ModRM.mod
 00: [Memory]
 01: 【メモリ + DISP8】
 の場合
 10: 【メモリ +
 DISP16】 の場合
 11: 登録

ModRM.reg Register code
 ModRM.rm If ModRM.mod = 11: register code
 000: [BX+SI].

001: [BX+DI].
 010: [BP+SI] の場合
 011: [BP+DI]
 100: [SI] の場合
 101: [DI] の場合
 110: [BP] または [DISP16] ModRM.mod=0
 のとき 111: [BX] となります。

Protected and Long modes

ModRM.mod
 00: [Memory]
 01: 【メモリ + DISP8】 の場合
 10: 【メモリ + DISP32】 の場合
 11: 登録

ModRM.reg Register code
 ModRM.rm If ModRM.mod = 11: register code
 REX. b=0 REX. b=1

000: [RAX] 000: [R8]
 001: [RCX] 001: [R9]
 010: [RDX] を使用し 010: [R10]
 ています。 011: [R11]
 011: [RBX] を使用し 100: [SIB] の場合
 ています。 101: [DISP32] の場合
 100: [SIB] の場合 110: [R14]
 101: [rbp][disp32]. 111: [R15]
 110: 【RSI】 の場合
 111: [RDI] の場合

ModRM.modフィールドは、Mod.rmフィールドと組み合わせてアドレッシングモードを決定します。例えば、mov ax, [0xa000]という命令は、（リアルモードでは）ModRM.mod = 0 (Memory)、ModRM.rm = 6 (DISP16) を使用します。ModRM.regにはAXのレジスタコードが入ります。mov ax, [bx+0xa000]を使用する場合、ModRM.rmは2 (Memory+DISP16)、ModRM.rmは7となります。アセンブラーは、0xa000がワードサイズの変位であることから、ここでは0xa000をDISP8ではなくDISP16として扱います。アドレスサイズオーバーライドプレフィックスを使用して、DISP8形式を選択することができます。

上記の表を見ると、間接アドレッシングに使えるレジスタはあまり多くないことが推測できます。例えば、[BP]は有効なアドレッシングモードではありませんが、アセンブラーはmov ax, [bp]のような命令でうまく組み立てることができます。これらのアセンブラーがよく使うトリックは、ModRM.mod = 1 (Memory+DISP8) と ModRM.rm = 6 (BP) を設定することです。つまり、アセンブラーはこれを[BP+DISP8]アドレッシングモードに変換し、ディスプレイスメントを0に設定します。そのため、mov ax, [bp]はmov ax, [bp+0]にアセンブルされます。

プロテクトモードとロングモードには、より多くの機能を提供する別のアドレッシングモード、[SIB]が導入されています。SIBアドレッシングモードは、ModRM.rmモードと組み合わせることができます。例えば、プロテクトモードでmov eax, [ebx+edi*2+0xa000]は、ModRM.mod = 2 (Memory+DISP32)、ModRM.rm = 4 (SIBバイト) と変換されます。SIBバイトは、この命令でedx+edi*2をどのように抽出するかを示すもので、次のセクションで説明します。一部の命令では、拡張オペコードフィールドを使用しています。拡張オペコードとは、命令を識別する際にプライマリ・オペコードと共に使用される識別子です。これは3ビットのフィールドで、これらの命令のMod R/M.regフィールドに格納されます。拡張オペコード・フィールドを使用する命令は、レジスタ・オペランドやメモリ・アドレッシング・モードを格納するために、ModRM.rmとModRM.modを使用する場合があります。

SIBフィールド

SIB (Scale Index Base) バイトは、Mod R/M.rm = 4で、CPUがプロテクトモードまたはロングモードの場合のみ、Mod R/Mバイトの後に続きます。このバイトは、IA32およびIA64アーキテクチャに追加のアドレッシングモードを提供します。SIBアドレッシングは、Mod R/Mアドレッシングと組み合わせることで、幅広いアドレッシングモードを実現することができます。

Scale	Index	Base	
7		0	

```

SIB.Scale 00: Factor 1
  01: ファクター2
  10: ファクター4
  11: ファクター8
SIB.Index 標準のレジスタコードを使用
  VSIBの場合、VRレジスタコードを使用
  REX.x=1の場合、64ビットのレジスタコードを使用
  REX.x=1でVSIBの場合、VRレジスタコードを使用
SIB.Base Uses standard register codes
  REX.b=1の場合、64ビットのレジスタコードを使用

```

SIBバイトのレジスタフィールドの名前にもかかわらず、技術的にはどのようなレジスタコードも使用することができます。例えば、SIB.Baseにインデックスレジスターを入れることができます。

SIBバイトとMod R/Mを再び組み合わせて、どのように動作するかを説明します。先ほどの例では、`mov eax, [ebx+edi*2+0xa000]` とします。この例では、CPUはプロジェクトモードで動作しているものとします。EAXは非メモリレジスタなので、Mod R/M.regに入ります。また、Mod R/M.mod = 2で[Memory+DISP32]を有効にし、Mod R/M.rm = 4で[SIBバイト]を有効にする必要があります。EBXはベースレジスタ、EDIはインデックスレジスタです。EBXのレジスタコードは3、EDIのレジスタコードは7です。これを用いて、SIB.index = 7、SIB.base = 3に設定します。スケールファクターの2はSIB.scaleに入ります。

これをまとめると、Mod R/Mのバイトは10 000 100のバイナリ、SIBのバイトは10 111 011のバイナリとなります。32ビットのディスプレイスメントが0xa000、オペレーションコードが0x89であることから、例題の命令を次のように翻訳することができます。

`0x89 0x84 0xbb 0x00 0xa0 0x00 0x00`

これは`mov eax, [ebx+edi*2+0xa000]`の正しい変換になります。読みやすくするために、演算コードは茶色、Mod R/MとSIBのバイトは赤色、ディスプレイスメントフィールドは黒色で表示しています。これは、命令の正確なフォーマットに従っていることに注意してください：最初に主オペコード、次にMod R/Mバイト、次にSIBバイト、そして変位バイトが続きます。

上の命令の変位バイトが奇妙に見える場合は、IA32およびIA64アーキテクチャがリトルエンディアンであることを考慮してください。

変位分野

ディスプレイスメントフィールドは、Mod R/M.modがモード1(Memory+DISP8)またはモード2(Memory+DISP16またはMemory+DISP32)の場合のみ有効です。ディスプレイスメントは、バイト、ワード、または、ワード値で、Mod R/MやSIBバイトと組み合わせて、アドレッシングモードにディスプレイスメントを追加するために使用されます。ディスプレイスメントフィールドは、常にMod R/MまたはSIBバイトの後に続きます。

直前のフィールド

即値フィールドは、命令がオペランドとして要求する場合にのみ有効です。命令は8、16、または32ビットの即値を必要とする場合があります。このフィールドは、命令の最後のフィールドとして存在する必要があります。16ビットと32ビットの両方の値を許容する命令では、オペランドオーバーライドサイズプレフィックスが存在するかどうかによって、このフィールドのサイズを決定します。

Instruction tables

あるソフトウェアでは、命令の機械語翻訳を容易にするために、命令ルックアップテーブルが利用されています。このテーブルは、すべての命令のオペレーションコードとオペランドタイプを提供する汎用命令表を反映したものです。IA32のマニュアルやオンラインリファレンスなどを参考にして、テーブルを作成したり、機械語命令の作成に役立てたりします。これらのテーブルの設計はかなり異なります。これらのルックアップテーブルの読み方については、ドキュメントを読むことが重要です。

テーブルには、以下のような形で指示が表示されます。
`| 0x10 | ADC | R/MEM8 | R8 |` となっています。

+-----+-----+-----+-----+

これは、オペレーションコードが0x10のADC命令を表しています。R/MEM8が第1オペランド、R8が第2オペランドとなります。オペランドはテーブルのデザインによってさまざまな方法で表現されます。また、このテーブルには、命令セットに影響を与えるフラグ、命令が使用するオペコードバイトの形式（オペコードフィールドにレジスタIDを格納する場合など）、サポートされるプロセッサなどの追加情報が提示されることもあります。これらのテーブルは、サイズが非常に大きくなることがあります、いざれも通常上記の形式で提示される基本的な情報を提供しています。

上記の「R/MEM8」は、第1オペランドが「レジスタ」または「8ビットのメモリオレーション」であることを意味しています。R8は、第2オペランドが8ビットのレジスタであることを意味しています。命令がメモリオペランドを持つ場合は、Mod R/M（場合によってはSIBバイトも）が続く必要があります。また、命令が2つのレジスタオペランドを取る場合は、Mod R/Mバイトが続く必要があります。Mod R/Mバイトは、Mod R/M.rmとMod R/M.regのメモリアドレスモード情報または両方のレジスタコードを格納します。CPUは、オペコードにより、レジスタコードが8ビットレジスタであることを認識します。オペコードは、実行する命令だけでなく、その命令がどのようなオペランドを必要とするのかをCPUに伝えます。1つの命令は異なるタイプのオペランドを使用することができ、そのため同じ命令が複数のオペコードを占めることができます。例えば、上記の命令形式ではオペコード0x10を使用しています。その他の形式としては、以下のようないがありますが、これらに限定されるものではありません。

0x11 ADC R/MEM16/MEM32 R16/REG32
+-----+-----+-----+-----+
0x12 ADC R8 R/REG16/REG32
+-----+-----+-----+-----+
0x13 ADC R16/REG32 R/MEM16/MEM32
+-----+-----+-----+-----+

命令がREG16/REG32のようなオペランドを使用する場合、オペランドサイズオーバーライド接頭辞があるかどうかと、現在のCPUの動作モード（プロジェクトモードで動作しているか、リアルモードで動作しているか、ロングモードで動作しているか、など）に基づいて、使用するオペランドを推測する必要があります。例えば、`ADC ax, word ptr [0]`という命令をプロジェクトモードで実行している場合（アセンブリ言語の用語で言うとbits32やuse32命令を使用している場合）、この命令にはオペコード0x13を使用することができます。この命令は「ADC REG16, MEM16/MEM32」という形式であることがわかります。AXは第1オペランドで、

16ビットのレジスタ（REG16）です。しかし、[0]とは何でしょうか？それを知るためには、アドレスサイズのオーバーライドがないことと、プロジェクトモードであることを考慮します。アドレスサイズオーバーライドがないため、ネイティブサイズである32ビットのメモリアドレッシングを使用することになります。（詳細は、アドレスサイズオーバーライドのプレフィックスのセクションを参照してください）。これにより、ADC REG16, MEM32形式を選択することになります。（アドレス・サイズ・オーバーライドが存在する場合は、ADC REG16, MEM16形式を選択することになります）。

1つの命令が1つのレジスタ・オペランドしか持たない場合、オペランドがOPCode.regフィールドに格納されているかどうかを確認します。（一部の命令はスペースを節約するためにこのようにしております、これがシングルバイト命令を可能にしています。また、オペランドをOPCode.regとMod R/M.regまたはMod R/M.rmに格納するために2つのレジスタを利用する命令もあります）。

この例を完成させるために、次のことを記します。OPCode 0x13、AXレジスタコード(0)、アドレッシングモードは[Memory]でディスプレイスメントは0。isがプロテクトモードであるため、32ビットのMod R/Mフォームを使用します。Mod R/M.reg = 0 (selecting AX)、Mod R/M.rm = 5 (DISP32)、Mod R/M.mod = 0 ([MEMORY])です。これにより、Mod R/Mの値は00 000 101のバイナリになります。SIB]モードを使用しないため、SIBバイトを使用する必要はありません。(SIBを使用する例としては、`mov eax, [ebx+edi*2+0xa000]`の逆アセンブルの例を参照してください)。これらの情報をもとに、マシンコードを作成します。

`0x66 0x13 0x05 0x00 0x00 0x00 0x00`

OPCodeは茶色、Mod R/Mバイトは赤色で表示されています。ディスプレイスメントバイトはDISP32(Mod R/M.rmのため)なので、dwordでなければなりません。これは黒で識別されます。0x66はオペランドサイズオーバーライドのプレフィックスで、青で表示されています。REG32およびMEM32形式ではなく、REG16およびMEM16形式を選択するために、オペランドサイズオーバーライドのプレフィックスを使用します。プレフィックスを省略した場合は、以下のようになります。

`0x13 0x05 0x00 0x00 0x00 0x00`

プロテクトモードでは、これは`adc eax, dword ptr [0]`命令であり、望んでいたものではありませんでした。詳しくは、オペランドサイズオーバーライドの接頭辞の項をご覧ください。

`ADC ax, word ptr [0]`を REP ADC ax, word ptr ES:[0] にしたい場合は、ES オーバーライドプレフィックスと REP プレフィックスを使います。

`0xf3 0x26 0x66 0x13 0x05 0x00 0x00 0x00`

プレフィックスバイトの順序は重要ではありません。

Resources

The following resources are presented for supplemental reading. Please note that we do not provide support for these resources.

<http://ref.x86asm.net/> IA32 and IA64 instruction table

<http://www.sandpile.org/> Instruction format tables

http://wiki.osdev.org/X86-64_Instruction_Encoding IA32 and IA64 Instruction encoding

Conclusion

本章では、IA32およびIA64アーキテクチャにおける機械語プログラミングと命令エンコーディングの概要を説明しました。本章では、デバッガやツールチェーンの開発を促進するために、新たな方法で資料を提供することを目的としています。また、本章は、特定の命令をエミュレートする際に、命令表を参考にすることができます。

ご質問やご意見がありましたらお聞かせください。

~Mike () です。

OS開発シリーズ編集部



Operating Systems Development Series オペレーティングシステム開発 - スキャンコード

by Mike, 2008

このシリーズは、オペレーティングシステムの開発を一から実演し、教えることを目的としています。

はじめに

これは、すべてのスキャンコードを一覧にした資料です。keybordコントローラには、3つの定義されたスキャンコードセットがあります。

オリジナルXTスキャンコードセット

スキャンコードセット

KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK
A	1E	9E		9	0A	8A		[1A	9A
B	30	B0		`	29	89		INSERT	E0, 52	E0, D2
C	2E	AE		-	0C	8C		HOME	E0, 47	E0, 97
D	20	A0		=	0D	8D		PG UP	E0, 49	E0, C9
E	12	92		¥	2B	AB		DELETE	E0, 53	E0, D3
F	21	A1		BKSP	0E	8E		END	E0, 4F	E0, CF
G	22	A2		SPACE	39	B9		PG DN	E0, 51	E0, D1
H	23	A3		TAB	0F	8F		U ARROW	E0, 48	E0, C8
I	17	97		CAPS	3A	BA		L ARROW	E0, 4B	E0, CB
J	24	A4		L SHFT	2A	AA		D ARROW	E0, 50	E0, D0
K	25	A5		L CTRL	1D	9D		R ARROW	E0, 4D	E0, CD
L	26	A6		L GUI	E0, 5B	E0, DB		NUM	45	C5
M	32	B2		L ALT	38	B8		KP /	E0, 35	E0, B5
N	31	B1		R SHFT	36	B6		KP *	37	B7
O	18	98		R CTRL	E0, 1D	E0, 9D		KP -	4A	CA
P	19	99		R GUI	E0, 5C	E0, DC		KP +	4E	CE
Q	10	90		R ALT	E0, 38	E0, B8		KP EN	E0, 1C	E0, 9C
R	13	93		APPS	E0, 5D	E0, DD		KP .	53	D3
S	1F	9F		ENTER	1C	9C		KP 0	52	D2
T	14	94		ESC	01	81		KP 1	4F	CF
U	16	96		F1	3B	BB		KP 2	50	D0
V	2F	AF		F2	3C	BC		KP 3	51	D1
W	11	91		F3	3D	BD		KP 4	4B	CB
X	2D	AD		F4	3E	BE		KP 5	4C	CC
Y	15	95		F5	3F	BF		KP 6	4D	CD
Z	2C	AC		F6	40	C0		KP 7	47	C7
0	0B	8B		F7	41	C1		KP 8	48	C8
1	02	82		F8	42	C2		KP 9	49	C9
2	03	83		F9	43	C3]	1B	9B
3	04	84		F10	44	C4		;	27	A7
4	05	85		F11	57	D7		'	28	A8
5	06	86		F12	58	D8		,	33	B3
6	07	87		PRNT SCRN	E0, 2A, E0, 37	E0, B7, E0, AA		.	34	B4

7	08	88		SCROLL	46	C6		/	35	B5
8	09	89		PAUSE	E1, 1D, 45 E1, 9D, C5	-NONE-				

ACPIスキャンコード

Key	Make Code	Break Code
Power	E0, 5E	E0, DE
Sleep	E0, 5F	E0, DF
Wake	E0, 63	E0, E3

Windowsマルチメディアスキャンコード

Key	Make Code	Break Code
Next Track	E0, 19	E0, 99
Previous Track	E0, 10	E0, 90
Stop	E0, 24	E0, A4
Play/Pause	E0, 22	E0, A2
Mute	E0, 20	E0, A0
Volume Up	E0, 30	E0, B0
Volume Down	E0, 2E	E0, AE
Media Select	E0, 6D	E0, ED
E-Mail	E0, 6C	E0, EC
Calculator	E0, 21	E0, A1
My Computer	E0, 6B	E0, EB
WWW Search	E0, 65	E0, E5
WWW Home	E0, 32	E0, B2
WWW Back	E0, 6A	E0, EA
WWW Forward	E0, 69	E0, E9
WWW Stop	E0, 68	E0, E8
WWW Refresh	E0, 67	E0, E7
WWW Favorites	E0, 66	E0, E6

最近のキーボードのデフォルツスキャンコードセット

スキャンコードセット

KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK
A	1C	F0, 1C		9	46	F0, 46		[54	F0, 54
B	32	F0, 32		`	0E	F0, 0E		INSERT	E0, 70	E0, F0, 70
C	21	F0, 21		-	4E	F0, 4E		HOME	E0, 6C	E0, F0, 6C
D	23	F0, 23		=	55	F0, 55		PG UP	E0, 7D	E0, F0, 7D
E	24	F0, 24		¥	5D	F0, 5D		DELETE	E0, 71	E0, F0, 71
F	2B	F0, 2B		BKSP	66	F0, 66		END	E0, 69	E0, F0, 69
G	34	F0, 34		SPACE	29	F0, 29		PG DN	E0, 7A	E0, F0, 7A
H	33	F0, 33		TAB	0D	F0, 0D		U ARROW	E0, 75	E0, F0, 75
I	43	F0, 43		CAPS	58	F0, 58		L ARROW	E0, 6B	E0, F0, 6B
J	3B	F0, 3B		L SHFT	12	F0, 12		D ARROW	E0, 72	E0, F0, 72
K	42	F0, 42		L CTRL	14	F0, 14		R ARROW	E0, 74	E0, F0, 74
L	4B	F0, 4B		L GUI	E0, 1F	E0, F0, 1F		NUM	77	F0, 77

M	3A	F0, 3A		L ALT	11	F0, 11		KP /	E0, 4A	E0, F0, 4A
N	31	F0, 31		R SHFT	59	F0, 59		KP *	7C	F0, 7C
O	44	F0, 44		R CTRL	E0, 14	E0, F0, 14		KP -	7B	F0, 7B
P	4D	F0, 4D		R GUI	E0, 27	E0, F0, 27		KP +	79	F0, 79
Q	15	F0, 15		R ALT	E0, 11	E0, F0, 11		KP EN	E0, 5A	E0, F0, 5A
R	2D	F0, 2D		APPS	E0, 2F	E0, F0, 2F		KP .	71	F0, 71
S	1B	F0, 1B		ENTER	5A	F0, 5A		KP 0	70	F0, 70
T	2C	F0, 2C		ESC	76	F0, 76		KP 1	69	F0, 69
U	3C	F0, 3C		F1	05	F0, 05		KP 2	72	F0, 72
V	2A	F0, 2A		F2	06	F0, 06		KP 3	7A	F0, 7A
W	1D	F0, 1D		F3	04	F0, 04		KP 4	6B	F0, 6B
X	22	F0, 22		F4	0C	F0, 0C		KP 5	73	F0, 73
Y	35	F0, 35		F5	03	F0, 03		KP 6	74	F0, 74
Z	1A	F0, 1A		F6	0B	F0, 0B		KP 7	6C	F0, 6C
0	45	F0, 45		F7	83	F0, 83		KP 8	75	F0, 75
1	16	F0, 16		F8	0A	F0, 0A		KP 9	7D	F0, 7D
2	1E	F0, 1E		F9	01	F0, 01]	5B	F0, 5B
3	26	F0, 26		F10	09	F0, 09		;	4C	F0, 4C
4	25	F0, 25		F11	78	F0, 78		'	52	F0, 52
5	2E	F0, 2E		F12	07	F0, 07		,	41	F0, 41
6	36	F0, 36		PRNT SCRN	E0, 12, E0, 7C	E0, F0, 7C, E0, F0, 12		.	49	F0, 49
7	3D	F0, 3D		SCROLL	7E	F0, 7E		/	4A	F0, 4A
8	3E	F0, 3E		PAUSE	E1, 14, 77, E1, F0, 14, F0, 77	-NONE-				

ACPIスキャンコード

Key	Make Code	Break Code
Power	E0, 37	E0, F0, 37
Sleep	E0, 3F	E0, F0, 3F
Wake	E0, 5E	E0, F0, 5E

Windowsマルチメディアスキャンコード

Key	Make Code	Break Code
Next Track	E0, 4D	E0, F0, 4D
Previous Track	E0, 15	E0, F0, 15
Stop	E0, 3B	E0, F0, 3B
Play/Pause	E0, 34	E0, F0, 34
Mute	E0, 23	E0, F0, 23
Volume Up	E0, 32	E0, F0, 32
Volume Down	E0, 21	E0, F0, 21
Media Select	E0, 50	E0, F0, 50
E-Mail	E0, 48	E0, F0, 48
Calculator	E0, 2B	E0, F0, 2B
My Computer	E0, 40	E0, F0, 40
WWW Search	E0, 10	E0, F0, 10
WWW Home	E0, 3A	E0, F0, 3A
WWW Back	E0, 38	E0, F0, 38

WWW Forward	E0, 30	E0, F0, 30
WWW Stop	E0, 28	E0, F0, 28
WWW Refresh	E0, 20	E0, F0, 20
WWW Favorites	E0, 18	E0, F0, 18

ATマザーボード用PS/2スキャンコードセット

KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK	-----	KEY	MAKE	BREAK
A	1C	F0, 1C		9	46	F0, 46		[54	F0, 54
B	32	F0, 32		`	0E	F0, 0E		INSERT	67	F0, 67
C	21	F0, 21		-	4E	F0, 4E		HOME	6E	F0, 6E
D	23	F0, 23		=	55	F0, 55		PG UP	6F	F0, 6F
E	24	F0, 24		¥	5C	F0, 5C		DELETE	64	F0, 64
F	2B	F0, 2B		BKSP	66	F0, 66		END	65	F0, 65
G	34	F0, 34		SPACE	29	F0, 29		PG DN	6D	F0, 6D
H	33	F0, 33		TAB	0D	F0, 0D		U ARROW	63	F0, 63
I	43	F0, 48		CAPS	14	F0, 14		L ARROW	61	F0, 61
J	3B	F0, 3B		L SHFT	12	F0, 12		D ARROW	60	F0, 60
K	42	F0, 42		L CTRL	11	F0, 11		R ARROW	6A	F0, 6A
L	4B	F0, 4B		L WIN	8B	F0, 8B		NUM	76	F0, 76
M	3A	F0, 3A		L ALT	19	F0, 19		KP /	4A	F0, 4A
N	31	F0, 31		R SHFT	59	F0, 59		KP *	7E	F0, 7E
O	44	F0, 44		R CTRL	58	F0, 58		KP -	4E	F0, 4E
P	4D	F0, 4D		R WIN	8C	F0, 8C		KP +	7C	F0, 7C
Q	15	F0, 15		R ALT	39	F0, 39		KP EN	79	F0, 79
R	2D	F0, 2D		APPS	8D	F0, 8D		KP .	71	F0, 71
S	1B	F0, 1B		ENTER	5A	F0, 5A		KP 0	70	F0, 70
T	2C	F0, 2C		ESC	08	F0, 08		KP 1	69	F0, 69
U	3C	F0, 3C		F1	07	F0, 07		KP 2	72	F0, 72
V	2A	F0, 2A		F2	0F	F0, 0F		KP 3	7A	F0, 7A
W	1D	F0, 1D		F3	17	F0, 17		KP 4	6B	F0, 6B
X	22	F0, 22		F4	1F	F0, 1F		KP 5	73	F0, 73
Y	35	F0, 35		F5	27	F0, 27		KP 6	74	F0, 74
Z	1A	F0, 1A		F6	2F	F0, 2F		KP 7	6C	F0, 6C
0	45	F0, 45		F7	37	F0, 37		KP 8	75	F0, 75
1	16	F0, 16		F8	3F	F0, 3F		KP 9	7D	F0, 7D
2	1E	F0, 1E		F9	47	F0, 47]	5B	F0, 5B
3	26	F0, 26		F10	4F	F0, 4F		;	4C	F0, 4C
4	25	F0, 25		F11	56	F0, 56		'	52	F0, 52
5	2E	F0, 2E		F12	5E	F0, 5E		,	41	F0, 41
6	36	F0, 36		PRNT SCRN	57	F0, 57		.	49	F0, 49
	3D	F0, 3D		SCROLL	5F	F0, 5F		/	4A	F0, 4A
8	3E	F0, 3E		PAUSE	62	F0, 62				

Conclusion

Until next time,

~Mike ();

BrokenThorn Entertainment. Currently developing EvolutionEngine and MicroOS Operating System.

Questions or comments? Feel free to [Contact me](#).