

特 別 研 究 報 告 書

題 目

Mint オペレーティングシステムを用いたNICドライ
バの割り込みデバッグ手法の実現

指導教員

報 告 者

藤田 将輝

岡山大学工学部 情報系学科

平成 27 年 2 月 6 日 提出

要約

OS のデバッグ手法として VM を用いたものがある．これは VM 上で 2 つの OS を走行させ，一方の OS から他方の OS へ任意のタイミングで割り込みを発生させることにより，バグを再現し，デバッグを支援するものである．しかし，VM を用いたデバッグ手法では，VM とハイパーバイザ間の処理の遷移に伴う処理負荷が発生する．このため，一定間隔で発生する割り込みや短い間隔で発生するバグのように処理負荷が影響する割り込み処理のデバッグが困難である．

そこで，Mint を用いた OS のデバッグ手法が提案されている．Mint は 1 台の計算機上で複数の OS が計算機資源を分割占有して走行できる．本研究では，Mint において，NIC による割り込みを任意に挿入できる環境を実現する．具体的には，割り込み元 OS から NIC ドライバが割り込み処理をするパケットを共有メモリに格納し，割り込み元 OS が占有しているコアから割り込み先の OS が占有しているコアへ IPI を送信することにより，割り込みハンドラが動作し，NIC ドライバが共有メモリからパケットを取得することで割り込み処理を発生させるものである．これにより，NIC ドライバの割り込みにより発生するバグを再現し，デバッグを支援することができる．

目次

1	はじめに	1
2	関連研究	2
2.1	仮想計算機を用いたデバッグ支援機構	2
2.2	概要と処理流れ	2
2.2.1	割り込み挿入法の概要	2
2.2.2	割り込み挿入法の処理流れ	2
2.2.3	ロギング/リプレイ手法の概要	4
2.2.4	ロギングの処理流れ	4
2.2.5	リプレイの処理流れ	5
2.2.6	問題点	6
3	Mint オペレーティングシステム	8
3.1	Mint の設計方針	8
3.2	Mint の構成	8
3.3	Mint を用いたデバッグ支援環境	9
3.3.1	目的	9
3.3.2	Mint を用いたデバッグ支援環境の概要	9
3.3.3	Mint を用いたデバッグ支援環境の処理流れ	11
4	NIC ドライバの割り込みデバッグ環境の設計	12
4.1	NIC ドライバに適用することの目的	12
4.2	NIC ドライバのパケット受信の流れ	12
4.3	NIC ドライバのデバッグ支援環境の概要	14
4.4	設計方針	15
4.5	設計にあたっての課題	15

4.6	課題への対処	16
4.7	対処を実現するために必要な機能	16
5	実装	17
5.1	NIC ドライバのデバッグ支援環境の処理流れ	17
5.2	必要な機能の実現	18
5.2.1	割り込みジェネレータ	18
5.2.2	共有メモリへの格納	19
5.2.3	受信バッファ状態の更新	19
5.2.4	IPI の送信	19
5.2.5	共有メモリからパケットを取得する割り込みハンドラ	19
5.2.6	割り込みハンドラの登録	20
6	評価	21
6.1	目的	21
6.2	方法	21
7	おわりに	22

目 次

2.1	割り込み挿入法の処理流れ	3
2.2	ロギングの処理流れ	4
2.3	リプレイの処理流れ	5
3.1	Mint の構成	9
3.2	Mint を用いたデバッグ支援環境の概要	10
3.3	Mint を用いたデバッグ支援環境の処理流れ	11
4.1	NIC ドライバのパケット受信処理流れ	13
4.2	NIC ドライバのデバッグ支援環境の概要	14
5.1	NIC ドライバのデバッグ支援環境の処理流れ	18

表 目 次

第 1 章

はじめに

近年，OS の多機能化に伴って，OS のデバッグが重要視されている．OS のデバッグ，特に割り込み処理に関するデバッグは非同期的な処理であるため，デバッグが困難になっている．このデバッグを支援する方法として，仮想計算機を用いたものがある．仮想計算機を用いて，デバッグ支援 OS とデバッグ対象 OS の 2 つの OS を動作させ，デバッグ支援 OS がデバッグ対象 OS へ割り込みを挿入させたり，デバッグ支援 OS がデバッグ対象 OS の動作を再現したりすることで，バグを再現し，デバッグを支援する．しかし，仮想計算機を用いると，仮想計算機とハイパーバイザ間の処理の遷移に伴う処理負荷が存在するため，一定間隔で発生する割り込みや，短い間隔で発生するバグのように，処理負荷が影響する割り込み処理のデバッグが困難である．このため，Mint を用いたデバッグ手法が提案されている．Mint は仮想化を用いずに複数の Linux を動作できる OS である．このため，Mint を用いてデバッグ支援環境を構築すると，ハイパーバイザが存在しないため，処理の遷移に伴う処理負荷も同様に存在しなくなる．これにより，一定間隔で発生する割り込み割り込みや短い間隔で発生するバグのデバッグが可能になる．本研究では，非同期的な割り込みが頻繁に発生する NIC ドライバに Mint を用いたデバッグ支援環境を構築することで，NIC ドライバの割り込み処理のデバッグを支援する．

第 2 章

関連研究

2.1 仮想計算機を用いたデバッグ支援機構

OS のデバッグを支援する環境の既存研究として仮想計算機を用いたものがある．仮想計算機を用いたデバッグ支援環境は大きく分けて 2 つある．割り込み挿入法とロギング/リプレイ手法である．これらについて以下で説明する．

2.2 概要と処理流れ

2.2.1 割り込み挿入法の概要

割り込み挿入法はハイパーバイザ上で動作するデバッグ支援 OS とデバッグ対象 OS によって構成される．プログラマがデバッグ対象 OS の割り込みを挿入したいコード位置でハイパーコールを挿入する．デバッグ対象 OS の走行時に，ハイパーコールを挿入した位置で割り込みを発生させ，バグを再現し，デバッグを支援する．

2.2.2 割り込み挿入法の処理流れ

割り込み挿入法の処理流れについて図 2.1 に示し，以下で説明する．

(1) 割り込み発生要求

デバッグ対象 OS に挿入したハイパーコールにより，デバッグ対象 OS がハイパーバイザのデバッグ支援機構へ割り込み発生要求を行う．その後，デバッグ対象 OS の処理を中断し，ハイパーバイザへ処理が遷移する．

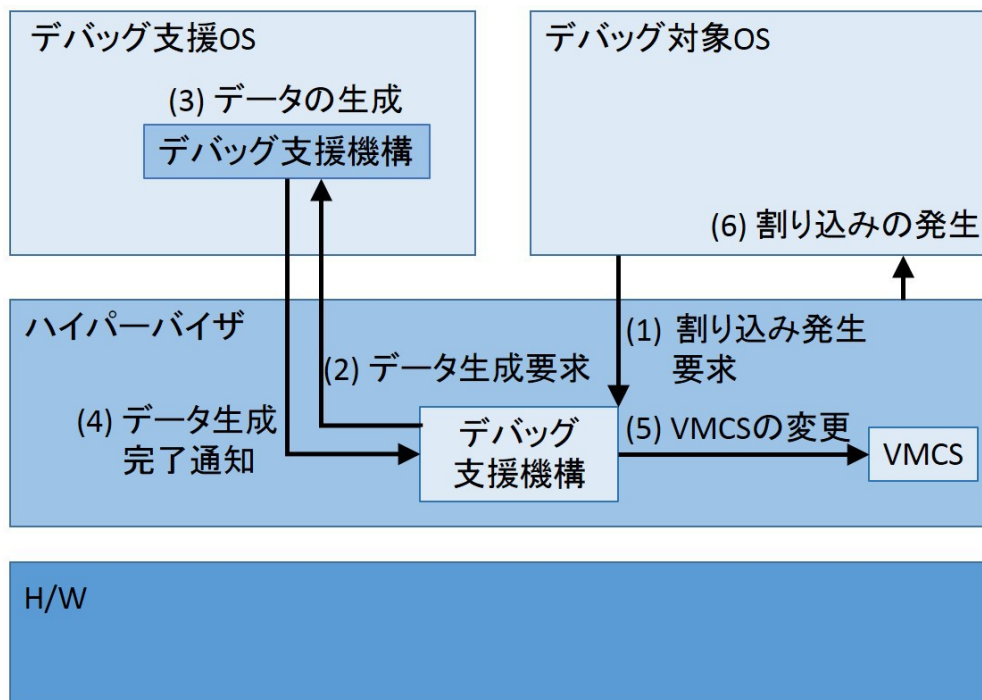


図 2.1 割り込み挿入法の処理流れ

(2) データ生成要求

ハイパーバイザのデバッグ支援機構がデバッグ支援 OS のデバッグ支援機構へ割り込みに必要なデータの生成要求を行う。割り込みに必要なデータとは、パケットやキーコードである。

(3) データの生成

デバッグ支援 OS のデバッグ支援機構が割り込みに必要なデータを生成する。

(4) データ生成完了通知

デバッグ支援 OS のデバッグ支援機構がハイパーバイザのデバッグ支援機構へデータの生成完了を通知する。

(5) VMCS の変更

ハイパーバイザのデバッグ支援機構が VMCS の内容を変更する。これにより、処理がハイパーバイザからデバッグ対象 OS へ処理が遷移するとき割り込みが発生する。

(6) 割り込み発生

デバッグ対象 OS へ処理が遷移し、割り込みが発生する。

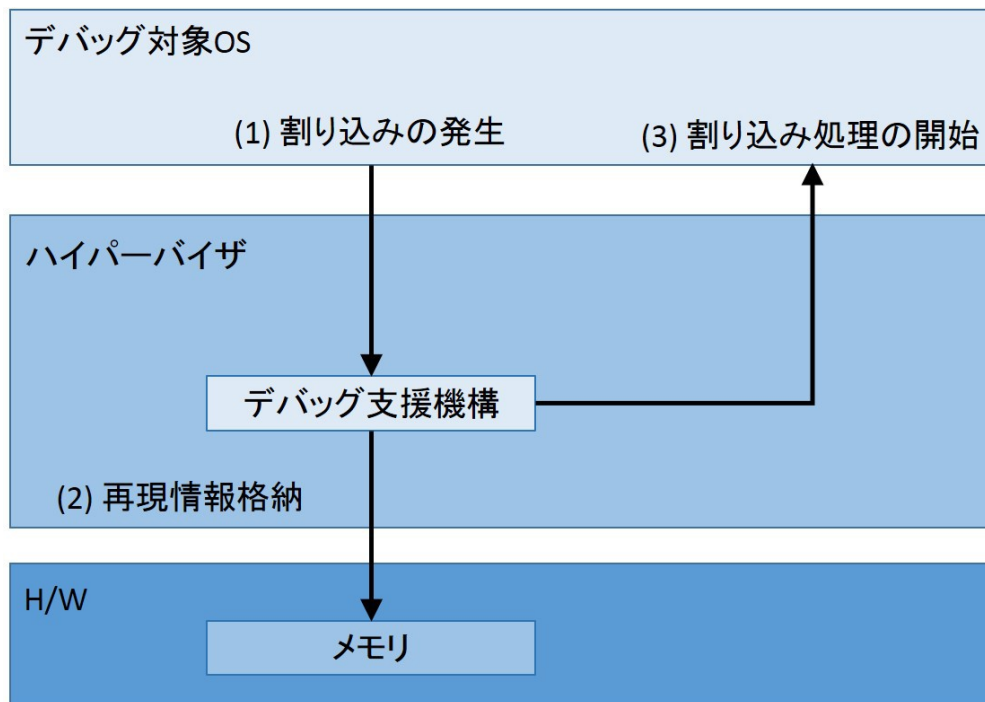


図 2.2 ログイングの処理流れ

2.2.3 ログイング/リプレイ手法の概要

ログイング/リプレイ手法はハイパーバイザまたはホスト OS 上で動作するデバッグ対象 OS により構成される。この手法はデバッグ対象 OS がバグを起こすまでの流れを保存し、再現することで、デバッグを支援する。ここでログイングとは OS の動作の流れを保存することで、リプレイとは保存した流れを再現することである。また、処理の流れを再現するための情報として、以下のような再現情報がある。

- (1) 割り込みの種類，割り込み発生アドレス，および分岐命令を経由した回数
割り込みの種類とこの割り込みが発生したアドレス，および分岐命令を経由した回数である。
- (2) 割り込み発生時に使用するデータ
キーコードや，パケットなどのような割り込み処理で扱うデータである。

2.2.4 ログイングの処理流れ

ログイングの処理流れについて図 2.2 に示し，以下で説明する。

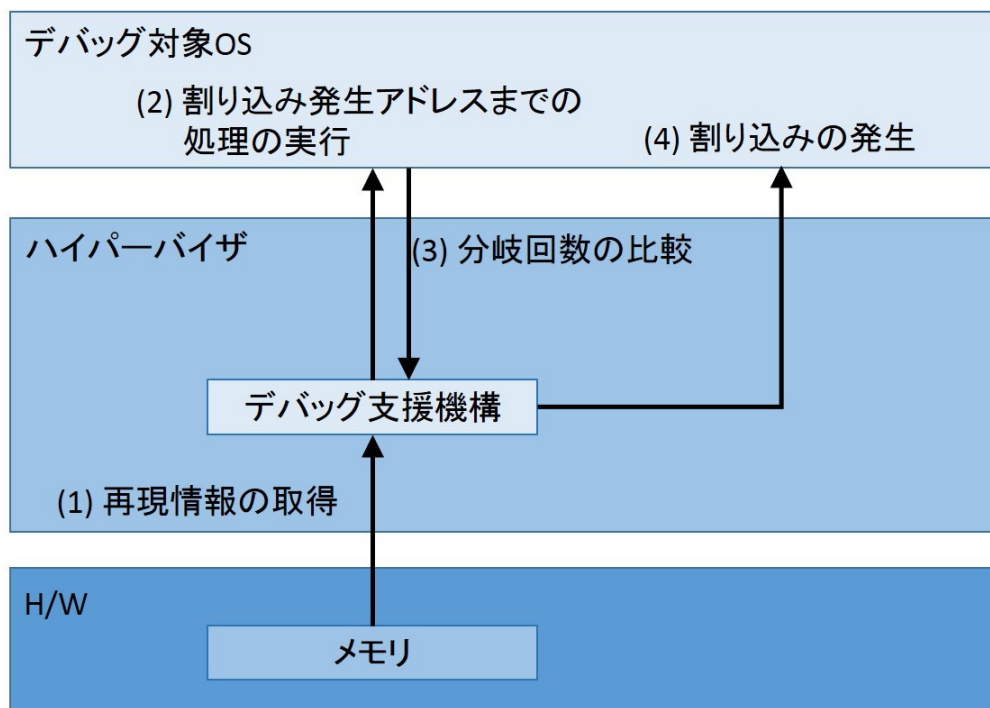


図 2.3 リプレイの処理流れ

(1) 割り込みの発生

デバッグ対象 OS に割り込みが発生すると、処理を中断し、ハイパーバイザに処理が遷移する。

(2) 再現情報の格納

ハイパーバイザのデバッグ支援機構が再現情報をメモリに格納する。

(3) 割り込み処理の開始

ハイパーバイザからデバッグ対象 OS へ処理が遷移し、デバッグ対象 OS が中断していた割り込み処理を再開する。

2.2.5 リプレイの処理流れ

リプレイの処理流れについて図 2.3 に示し、以下で説明する。

(1) 再現情報の取得

ハイパーバイザのデバッグ支援機構がメモリから再現情報を取得する。

(2) 割り込み発生アドレスまでの処理の実行

取得した再現情報よりデバッグ対象 OS が割り込みが発生するアドレスまで命令を実行する。

(3) 分岐回数の比較

ハイパーバイザのデバッグ支援機構が再現情報の分岐回数と、現在のデバッグ対象 OS の分岐回数を比較する。

(4) 割り込みの発生

デバッグ対象 OS へ割り込みが発生する。

2.2.6 問題点

割り込み挿入法の問題点について以下で説明する。

(1) 実計算機上で発生する間隔での複数割り込みの発生が困難

割り込み挿入法では割り込みを発生させる際、OS のコードの任意の位置にハイパーコールを挿入することで割り込みを発生させる。コードが実行されるタイミングは OS の処理速度に依存する。このため、CPU へ発生する間隔で複数の割り込みを発生させようとした際に、この間隔を調整するのはハイパーコールの間隔を調整することで行うが、これは非常に困難である。つまり、実計算機上で発生する間隔での複数の割り込み (以下、実割り込み) を発生させることが困難である。

また、ロギング/リプレイ手法の問題点について以下で説明する。

(2) 任意のタイミングでの割り込み発生が困難
ロギング/リプレイ手法は、ロギング時に発生した割り込みに対する処理をリプレイ時に確認できる。しかし、任意のタイミングで割り込みを発生させるためには、再現情報として割り込みを発生させるアドレスと分岐回数をプログラマが用意しなければならない。これらの指定が困難であるため、任意のタイミングで割り込みを発生させることが困難である。

(3) 実割り込みの発生が困難
ロギング/リプレイ手法は、ロギングにおけるデバッグ対象 OS とハイパーバイザの間の処理の遷移や再現情報の格納による処理負荷が発生する。このため、実割り込みがロギング中に発生しないと考えられる。ロギング中に実割り込みが発生しない場合、実割り込みを再現するための再現情報を保存できない。このため、実割り込みの発生が困難である。

これらの問題点から，割り込み処理のデバッグには，デバッグ対象 OS がデバッグ支援機構の処理負荷の影響を受けない環境が必要である．

第 3 章

Mint オペレーティングシステム

3.1 Mint の設計方針

Mint とは 1 台の計算機上で仮想化を用いずに計算機資源を分割することによって複数の Linux を動作させる方式である。Mint の設計方針として以下の 2 つが挙げられる。

- (1) 全ての Linux が相互に処理負荷の影響を与えない。
- (2) 全ての Linux が入出力性能を十分に利用できる。

3.2 Mint の構成

Mint では、1 台の計算機上で CPU、メモリ、およびデバイスを分割し、各 OS が占有する。Mint の構成例を図 3.1 に示し、説明する。Mint では、最初に起動する OS を OS ノード 0 とし、起動順に OS ノード 1、OS ノード 2、... とする。

- (1) CPU
コア単位で分割し、各 OS ノードがコアを 1 つ以上占有する。
- (2) メモリ空間分割し、各 OS ノードが分割領域を占有する。
- (3) デバイス
デバイス単位で分割し、各 OS ノードが指定されたデバイスを占有する。

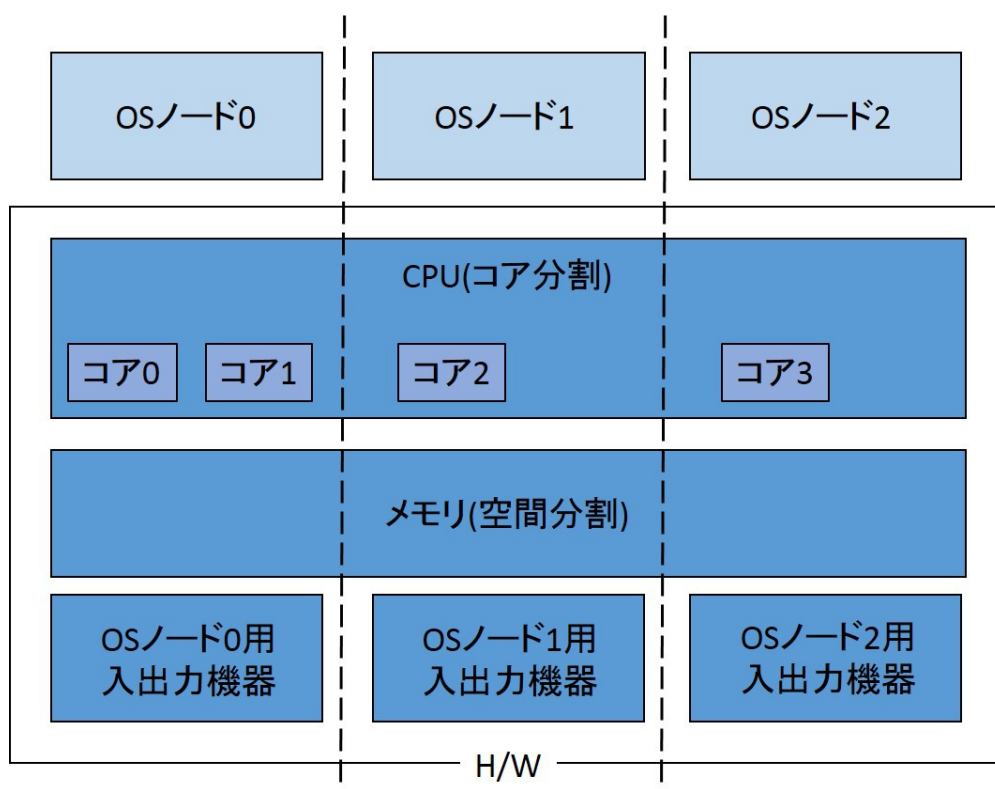


図 3.1 Mint の構成

3.3 Mint を用いたデバッグ支援環境

3.3.1 目的

前章で述べたように，割り込み処理のデバッグにはデバッグ対象 OS がデバッグ支援機構の処理負荷の影響を受けない環境が必要である．そこで，仮想化を用いずに複数の Linux を動作できる Mint を用いたデバッグ支援環境が提案されている．Mint を用いることで，実割り込みの発生が可能になる．また，任意のタイミングでの割り込み発生も可能になる．

3.3.2 Mint を用いたデバッグ支援環境の概要

Mint 用いたデバッグ支援環境の概要について図 3.2 に示し，説明する．Mint を用いたデバッグ支援環境はデバッグ支援 OS とデバッグ対象 OS の 2 つの OS が動作する．デバッグ支援 OS にはデバッグ支援機構があり，デバッグ対象 OS へ割り込みを発生させる．また，割り込みジェネレータをデバッグ支援 OS のアプリケーションとして実装している．これらの

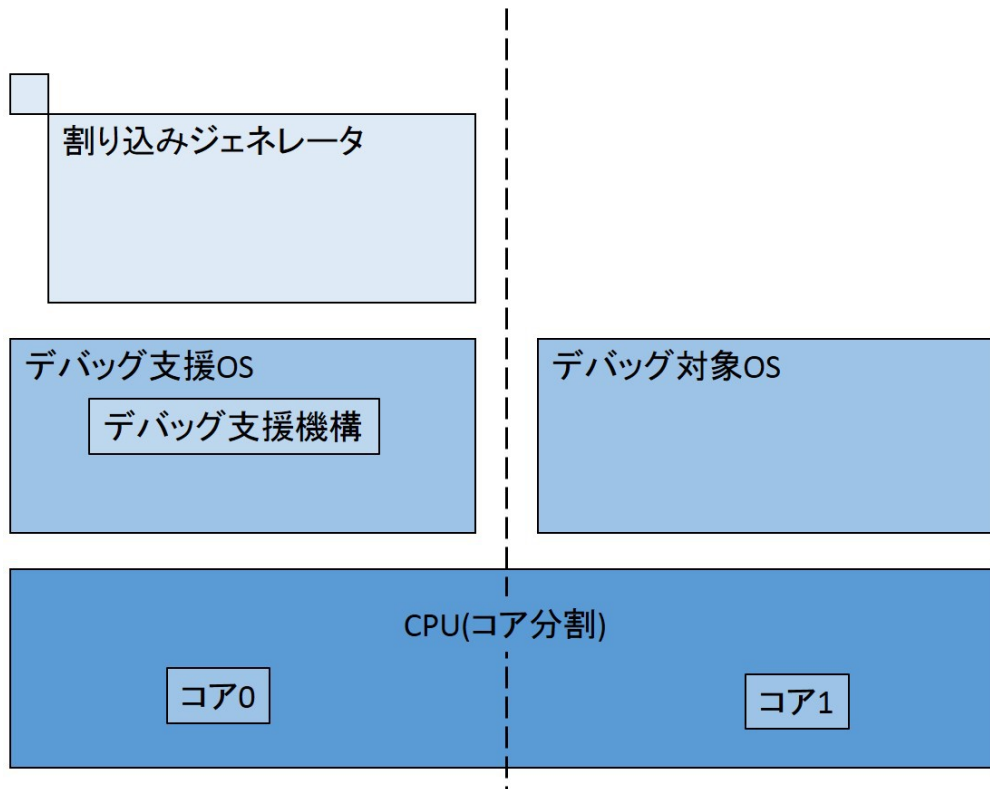


図 3.2 Mint を用いたデバッグ支援環境の概要

割り込みジェネレータ，デバッグ支援機構，デバッグ支援 OS，およびデバッグ対象 OS について以下で説明する．

(1) 割り込みジェネレータ

プログラマが割り込み情報を指定する際に利用するアプリケーションである．なお，割り込み情報とは割り込みの種類，発生間隔，および発生回数を合わせた情報である．

(2) デバッグ支援機構

割り込みジェネレータから通知される割り込み情報をもとに，デバッグ対象 OS へ IPI を送信する機構である．

(3) デバッグ支援 OS

デバッグ支援機構を実装して走行する OS である．ここでは，コア 0 を占有して走行する．

(4) デバッグ対象 OS

デバッグ対象となる OS である．ここでは，コア 1 を占有して走行する．

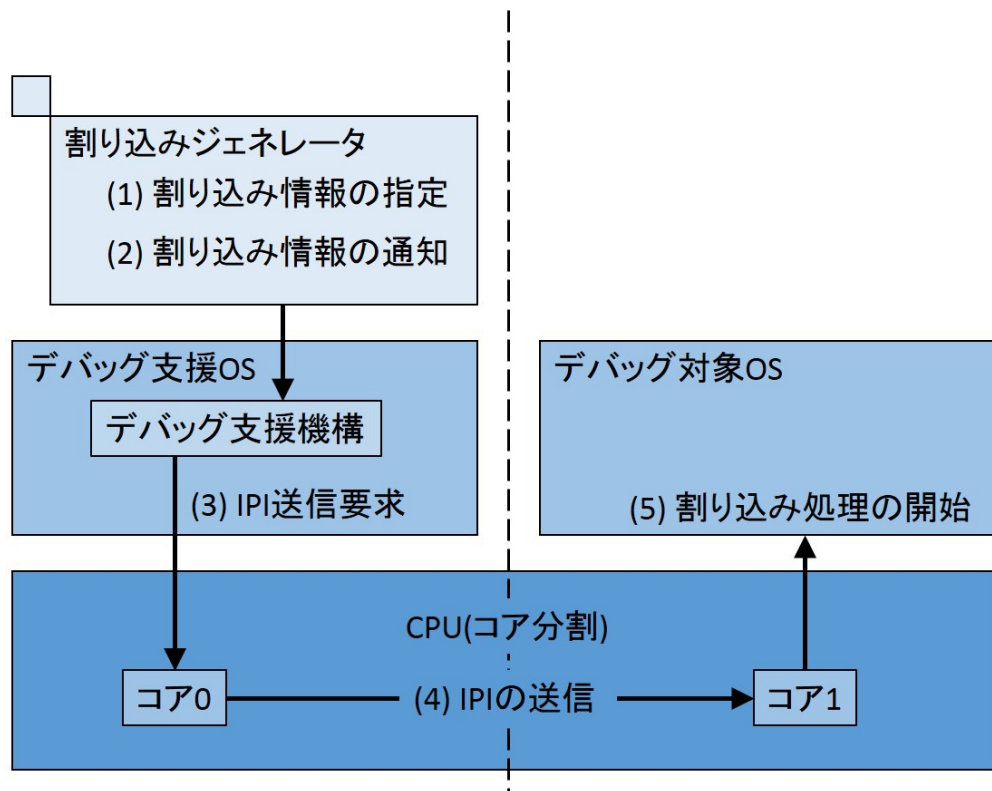


図 3.3 Mint を用いたデバッグ支援環境の処理流れ

3.3.3 Mint を用いたデバッグ支援環境の処理流れ

Mint を用いたデバッグ支援環境の処理流れを図 3.3 に示し、以下で説明する。

- (1) 割り込み情報の指定
デバッグ支援 OS 上で動作する AP でプログラマが割り込み情報を指定する。
- (2) 割り込み情報の通知
割り込み情報をデバッグ支援 OS のデバッグ支援環境に通知する。
- (3) IPI の送信要求
IPI の送信要求を行う。
- (4) IPI の送信
IPI を送信する。
- (5) 割り込み処理の開始
割り込み処理の開始。

第 4 章

NIC ドライバの割り込みデバッグ環境の設計

4.1 NIC ドライバに適用することの目的

割り込み処理におけるバグの原因として、非同期な割り込みがある。非同期な割り込みとは、割り込み先と割り込み元でタイミングを合わせないで発生する割り込みのことである。この非同期な割り込みが頻繁に発生するのがNIC ドライバである。NIC ドライバに Mint を用いた割り込みデバッグ支援環境を構築することでNIC ドライバのデバッグを支援する。また、本研究におけるデバッグ支援環境はNIC ドライバの受信処理に関するデバッグを行うものである。

4.2 NIC ドライバのパケット受信の流れ

NIC がパケットを受信してから NIC ドライバで割り込み処理が発生するまでの流れを図 4.1 に示し、以下で説明する。

(1) 受信バッファアドレスの取得

NIC が受信ディスクリプタから受信バッファのアドレスを取得する。

(2) 受信バッファへのパケットの格納

NIC が受信バッファへパケットを格納する。

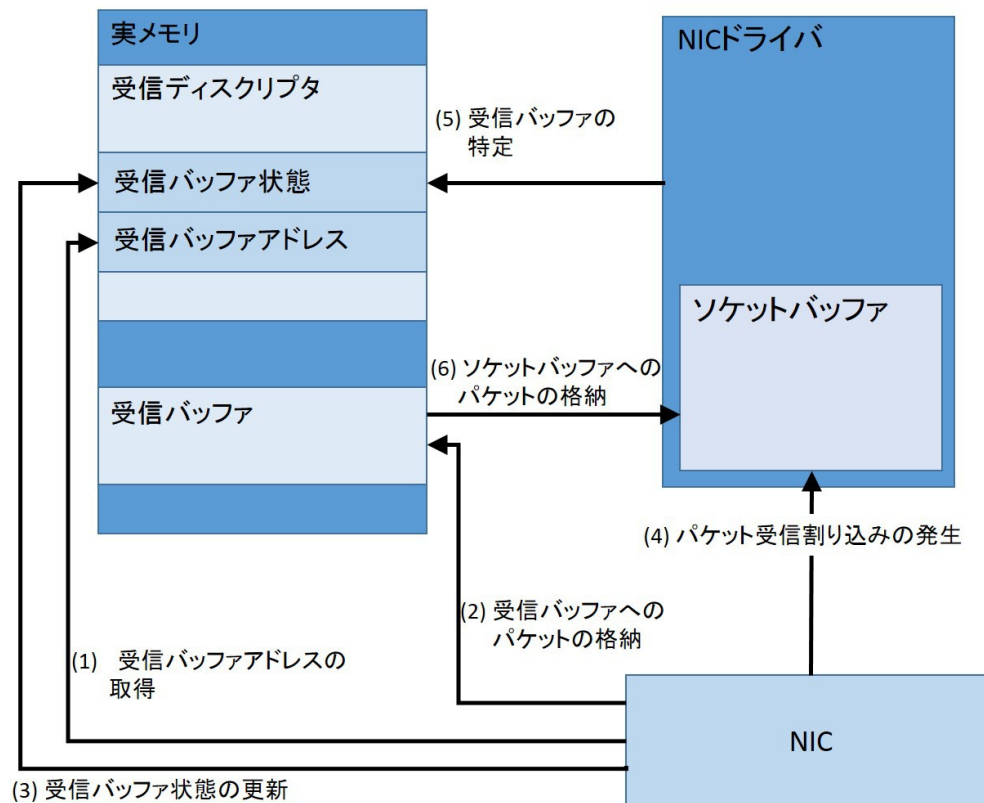


図 4.1 NIC ドライバのパケット受信処理流れ

(3) 受信バッファ状態の更新

NIC が受信ディスクリプタの受信バッファ状態を更新し、受信済み状態にする。

(4) パケット受信割り込みの発生

NIC から NIC ドライバに割り込みを発生する。

(5) 受信バッファの特定

NIC ドライバが受信ディスクリプタの受信バッファ状態を確認し、受信バッファを特定する。

(6) ソケットバッファへのパケットの格納

NIC ドライバが受信バッファからソケットバッファへパケットを取得し、格納する。

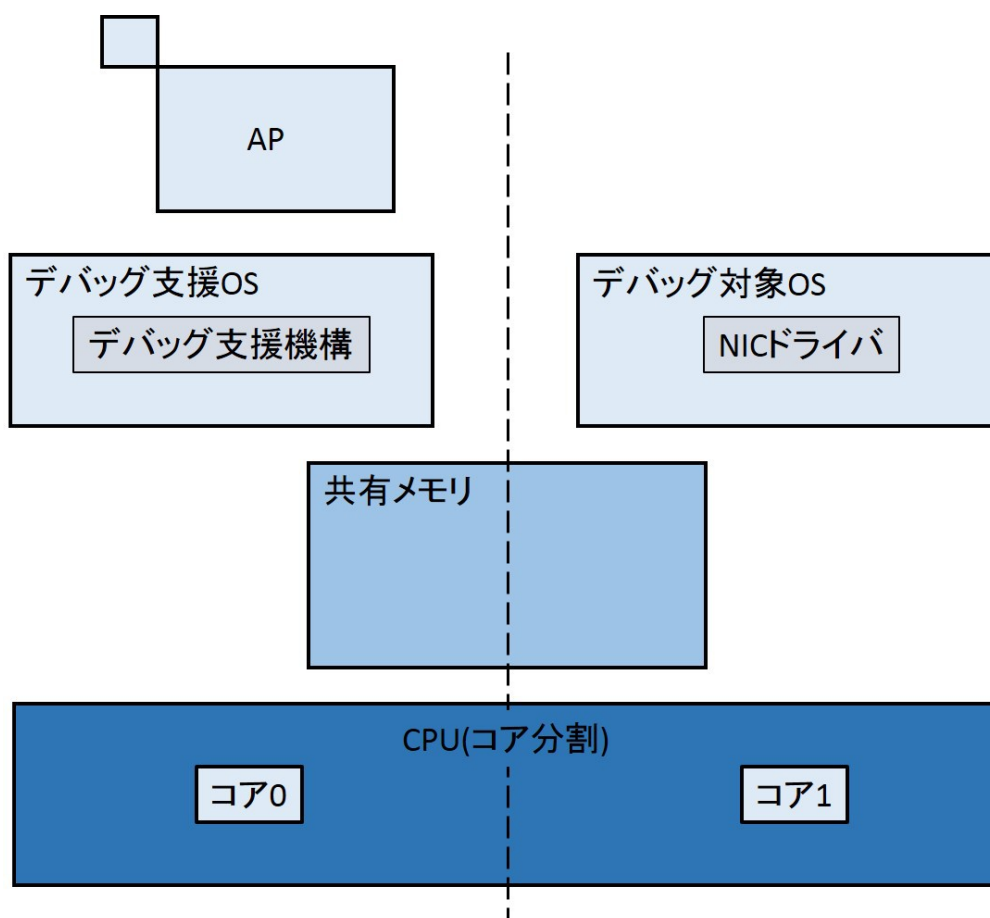


図 4.2 NIC ドライバのデバッグ支援環境の概要

4.3 NIC ドライバのデバッグ支援環境の概要

NIC ドライバのデバッグ支援環境の概要を図 4.2 に示す．NIC ドライバのデバッグ支援環境はデバッグ支援 OS，デバッグ対象 OS，共有メモリ，NIC ドライバ，デバッグ支援機構，CPU，およびアプリケーションによって構成される．それぞれについて以下で説明する．

(1) デバッグ支援 OS

NIC ドライバのデバッグを支援する OS である．この OS は NIC の動作を擬似する．

(2) デバッグ対象 OS

デバッグの対象となる OS である．本研究ではこの OS の NIC ドライバの受信処理のデバッグを対象としている．

(3) 共有メモリ

Mint の共有メモリである．これを用いてデバッグ支援 OS からデバッグ対象 OS の NIC ドライバへパケットを受け渡す．

(4) NIC ドライバ

デバッグ対象 OS 内で動作している．受信バッファを共有メモリとみなすように改変されている．また，割り込みの契機として IPI を用いる．

(5) デバッグ支援機構

ユーザが指定した情報に基づいてパケットを作成し，共有メモリに配置する．また，コア 0 へ IPI の送信要求を発信する．

(6) CPU

コアを分割してコア 0 をデバッグ支援 OS に，コア 1 をデバッグ対象 OS に占有させている．

(7) アプリケーション

ユーザがパケットの種類，割り込みの発生間隔を指定する．この情報からデバッグ支援機構によりパケットが生成される．

4.4 設計方針

Mint を用いた割り込みデバッグ環境を構築する際の設計方針を以下に示す．

- (1) デバッグ支援 OS が NIC に相当する動作を再現する．
- (2) NIC を用いずにデバッグ支援環境を構築する．
- (3) 任意のタイミングで割り込みを発生させられる．

4.5 設計にあたっての課題

設計にあたっての課題を以下に示す．

- (課題 1) 1 台の計算機上で動作する 2 つの Linux 間でのパケットの受け渡し．
- (課題 2) NIC を用いないため，割り込みの契機の変更．

(課題 3) 割り込みハンドラとして、受信バッファ相当のものからパケットを取得するものを登録する。

(課題 4) タイミングを調整できるようにする。

4.6 課題への対処

課題への対処を以下に示す。

(対処 1) 課題 1 への対処として Mint の共有メモリを用いて、パケットの受け渡しを行う。

(対処 2) 課題 2 への対処として、割り込みは IPI を用いて行う。

(対処 3) 課題 3 への対処として、NIC ドライバを改変することにより、共有メモリからパケットを取得し、処理をする割り込みハンドラを作成し、登録する。

(対処 4) 課題 4 の対処として、ユーザから割り込みの間隔を指定できるようにする。

4.7 対処を実現するために必要な機能

対処を実現するために必要な機能を以下に示す。

- (1) 割り込みジェネレータの作成
- (2) 割り込みジェネレータにおいての間隔の調整
- (3) 共有メモリへの格納
- (4) 受信バッファ状態の更新
- (5) IPI の送信
- (6) 共有メモリからパケットを取得する割り込みハンドラの作成
- (7) 割り込みハンドラの登録

第 5 章

実装

5.1 NIC ドライバのデバッグ支援環境の処理流れ

Mint を用いた NIC ドライバの割り込みデバッグ支援環境の処理流れを図 5.1 に示し、以下で説明する。

- (1) 割り込み情報の指定
割り込みジェネレータでユーザが割り込み情報を指定する。
- (2) 割り込み情報の通知
デバッグ支援 OS のデバッグ支援機構に指定した割り込み情報を通知する。
- (3) パケットの作成
デバッグ支援機構が割り込み情報からパケットを作成する。
- (4) 受信バッファへのパケットの格納
デバッグ支援機構が共有メモリの受信バッファへパケットを格納する。
- (5) 受信バッファ状態の更新
デバッグ支援機構が共有メモリの受信ディスクリプタの受信バッファ状態を更新する。
- (6) 割り込み発生要求
デバッグ支援機構がコア 0 へ IPI 送信要求を行う。
- (7) IPI の送信
コア 0 からコア 1 へ IPI を送信する。

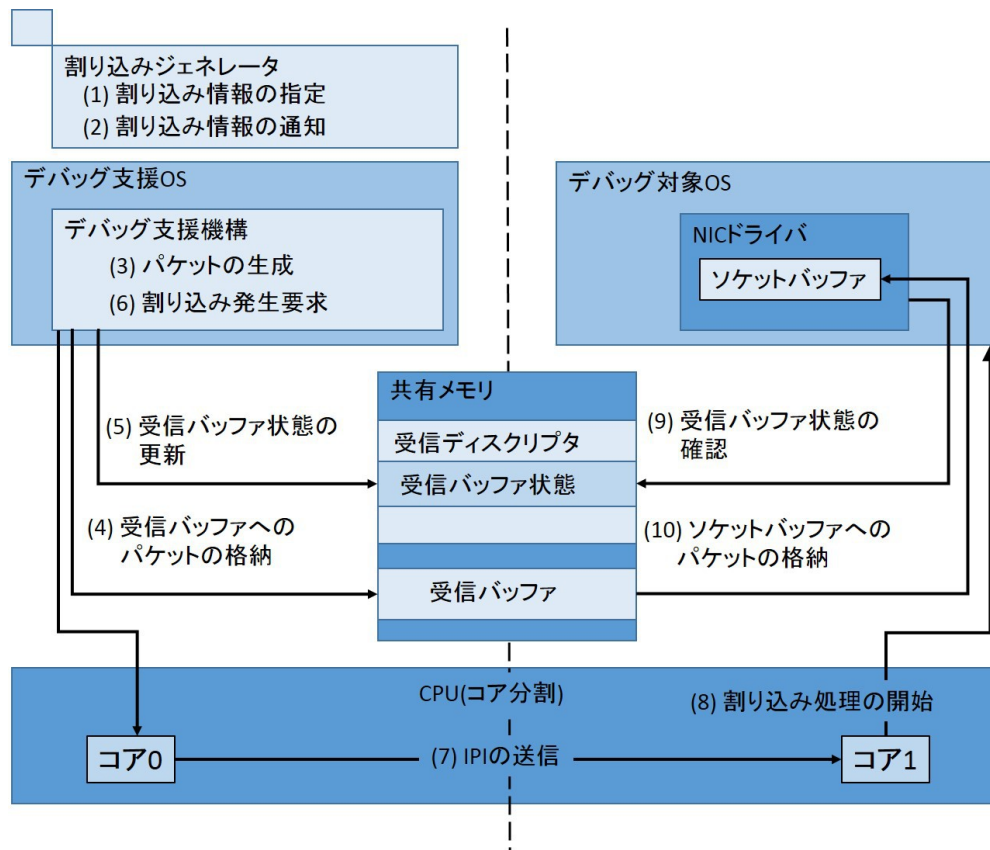


図 5.1 NIC ドライバのデバッグ支援環境の処理流れ

(8) 割り込み処理の開始

コア 1 が IPI を受信すると、デバッグ対象 OS の割り込みハンドラが動作する。

(9) 受信バッファ状態の確認

NIC ドライバが共有メモリの受信ディスクリプタ中の受信バッファ状態を確認する。

(10) ソケットバッファへのパケットの格納

NIC ドライバが共有メモリの受信バッファからパケットを取得し、ソケットバッファに格納する。

5.2 必要な機能の実現

5.2.1 割り込みジェネレータ

(1) 割り込みジェネレータの作成には NIC の送信用の関数を使う

(2) まだ未着手

5.2.2 共有メモリへの格納

Mint の共有メモリを利用してパケットの受け渡しを実現する。割り込みジェネレータによって作成されたパケットを共有メモリに配置する。これはシステムコールとして登録し、機能する。NIC ドライバの受信バッファはリングバッファとなっているため、パケットを配置するとカウンタをインクリメントし、これに対応する。また、共有メモリを NIC のドライバの受信バッファとみなすため、デバッグ対象 OS の NIC ドライバの初期化処理中において、受信バッファのアドレスを Mint の共有メモリのアドレスに変更する。

5.2.3 受信バッファ状態の更新

受信バッファは各要素ごとに受信バッファの情報を持っている。具体的には受信バッファのアドレス、受信バッファにパケットが格納されているか否か、という情報を持っている。NIC ドライバはこの情報を元にパケットを取得する。本来は NIC がパケットを受信バッファに格納する際にこれを変更する。本機構では NIC を使用しないため、共有メモリにパケットを格納した際にこの受信済み状態を更新する。

5.2.4 IPI の送信

IPI を送信する際の流れを以下に示す。

- (1) コアの持つ LAPIC(割り込みコントローラ) 中の ICR という IPI 送信用のレジスタに、ベクタ番号と LAPIC ID を書き込む。
- (2) LAPIC ID を参照し、この ID を持つコアへ IPI を送信する。
- (3) コアが IPI を受信すると、指定したベクタ番号に対応した割り込みハンドラが動作する。

5.2.5 共有メモリからパケットを取得する割り込みハンドラ

- (1) 共有メモリを受信バッファとみなし、受信バッファ状態を確認し、パケットを取得する。
- (2) ソケットバッファにパケットを格納し、上位層に送る。

5.2.6 割り込みハンドラの登録

- (1) NIC ドライバの初期化時に，作成した割り込みハンドラを登録する．

第 6 章

評価

6.1 目的

Mint を用いた NIC ドライバの割り込みデバッグ支援環境についての評価を行う．実際の NIC を用いた割り込み処理と本研究のデバッグ支援環境の割り込み処理を比べてどの程度の差があるのかを評価する．これにより，バグが発生した際と同程度の計算機の状態であることを示す．

6.2 方法

評価の方法について以下に示し，説明する．

(1) 割り込み間隔の評価

どれだけ短い間隔で連続の割り込みが発生させられるかを評価する．具体的には，デバッグ支援 OS からデバッグ対象 OS へ間隔を調整して割り込みを発生させ，どの程度の短さまで正常に割り込みが処理されるのかを評価するものである．

(2) CPU 負荷の評価

NIC を使った実際の割り込み処理と比べて，どの程度の CPU 負荷がかかっているのかを評価する．具体的には，NIC を用いた割り込み処理の CPU 負荷を調査し，本研究のデバッグ支援環境を用いた割り込み処理の CPU が実際の割り込み処理の CPU 負荷にどれだけ近付いているかを調査する．

第 7 章

おわりに

- (1) 本論文の流れを示す .
- (2) Mint を用いたデバッグ支援環境の有用性を示す .
- (3) 本研究では NIC ドライバの割り込み処理に , Mint を用いたデバッグ支援環境を適用させたことを示す .
- (4) 残された課題について述べる .