

# マイクロカーネル OS における NIC ドライバプロセス入替えの比較評価

## Comparative Evaluation of NIC Driver Process Replacement in Microkernel OS

澤田 淳<sup>†</sup>  
Jun Sawada

山内利宏<sup>†</sup>  
Toshihiro Yamauchi

谷口秀夫<sup>†</sup>  
Hideo Taniguchi

### 1. はじめに

サービスプログラムの変更において、サービス停止を抑制し、動作中のプログラムを変更できる機能が有効である。

モノリシックカーネル構造を有する Linux は、LKM (Loadable Kernel Module) のアンロードとロードによって OS プログラムの入替えを実現している。この入替えでは、使用中の LKM を入替えることができない。また、アンロード時にサービスが中断され、新たにロードされた LKM の内部状態に関する情報が初期化されるため、サービス再開に時間を要する。

本稿では、マイクロカーネル構造を有する MINIX3[1] と *AnT*[2] について、NIC ドライバプロセスの入替え時間とデータ送信時間を測定し、入替え処理がデータ送信処理に与える影響について比較評価する。

### 2. OS サーバ入替え機能

#### 2.1 MINIX3

MINIX3 の OS サーバの入替え処理は、OS サーバ入替え機能呼び出すプロセス (以降、入替え要求プロセス) により、全ての OS サーバの死活監視を行う Reincarnation Server (以降、RS サーバ) への入替え依頼を発行することで開始する。これを受け、RS サーバは、動作中の OS サーバ (以降、旧 OS サーバ) と同等のサービスを提供する新たな OS サーバ (以降、新 OS サーバ) を生成し、OS サーバの内部状態に関する情報 (以降、状態情報) の複製を保持する Data Store サーバ (以降、DS サーバ) に旧 OS サーバの状態情報を書き出す。その後、RS サーバは、新 OS サーバを起動し、新 OS サーバは、DS サーバから状態情報を読み込む。

ここで、RS サーバは、入替え要求を通常の処理依頼よりも優先して取得する。一方、DS サーバは、新旧 OS サーバの行う状態情報の読み込み/書き出し依頼を通常の処理依頼として受け取る。このため、DS サーバに対し、入替え要求とは無関係の依頼がある場合、入替え処理が待たされる。また、全ての OS サーバの状態情報を DS サーバが一元管理するため、DS サーバが不具合等により停止した場合、全ての OS サーバの状態情報を喪失する。

#### 2.2 *AnT*

*AnT* の OS サーバの入替え処理は、入替え要求プロセスにより、カーネルに入替え要求システムコールを発行することで開始する。これを受け、カーネルは、新 OS サーバを生成し、旧 OS サーバの保持する情報を新

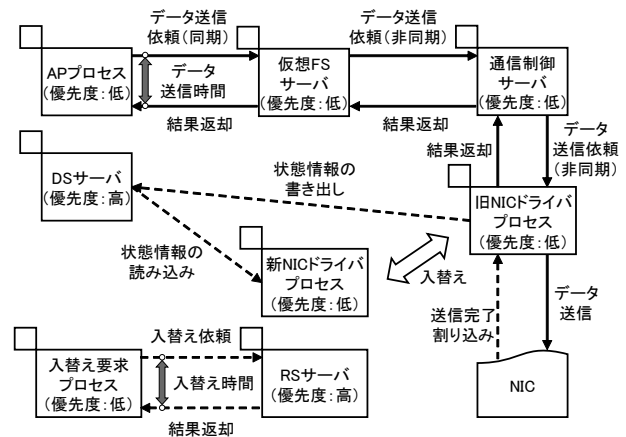


図 1 MINIX3 における NIC ドライバプロセス入替え

OS サーバに移譲する。なお、OS サーバの状態情報は、プロセス間の共有領域である ICA (Inter-core Communication Area) に格納されており、新旧 OS サーバ間で複写レスでの高速な引き継ぎが可能である。

また、ICA はプロセス空間から独立して存在するため、状態情報の操作が他のプロセスの処理依頼により遅延することや、特定の OS サーバの不具合により全ての OS サーバの状態情報を喪失することはない。

### 3. 評価

#### 3.1 評価内容

MINIX3 における NIC ドライバプロセス入替えの様子を図 1 に示す。入替え要求の発行から結果返却までの時間を入替え時間、AP プロセスにおけるデータ送信依頼の発行から結果返却までの時間をデータ送信時間とそれぞれ定義し、MINIX3 と *AnT* において、データ送信処理中に NIC ドライバプロセスを入替える場合の NIC ドライバプロセスの入替え時間とデータ送信時間を測定した。

送信側計算機、受信側計算機の 2 台ともに Intel<sup>®</sup> Core i7 (3.2GHz, 1 コアのみ使用) と Realtek 8139 PCI Ethernet Card (100Mbps) を搭載した計算機を用いて測定した。具体的には、1 個の AP プロセスから TCP/IP 通信による 1024byte のデータ送信処理と 1 ms の PU 処理を交互に 1000 回行う間のランダムな契機に NIC ドライバプロセスの入替えを 1 回行った。この測定を 1 回の試行とし、試行を 5 回繰り返した。

なお、*AnT* に RS サーバ、DS サーバ、および仮想 FS サーバは存在せず、プロセスの優先度は、入替え要求プロセスのみを高とし、これ以外を低とした。

<sup>†</sup> 岡山大学大学院自然科学研究科, Graduate School of Natural Science and Technology, Okayama University

表 1 入替え時間とデータ送信時間

	MINIX3				<i>AnT</i>			
	入替え時間 (ms)		データ送信時間 (ms)		入替え時間 (ms)		データ送信時間 (ms)	
	データ送信無しの場合	データ送信有りの場合	入替え無しの場合	入替え有りの場合	データ送信無しの場合	データ送信有りの場合	入替え無しの場合	入替え有りの場合
最大値	0.316	0.336	0.024	0.025	0.579	10.331	0.018	0.114
平均値	0.313	0.316	0.012	0.012	0.568	10.140	0.011	0.011
最小値	0.311	0.311	0.009	0.009	0.562	9.993	0.009	0.009

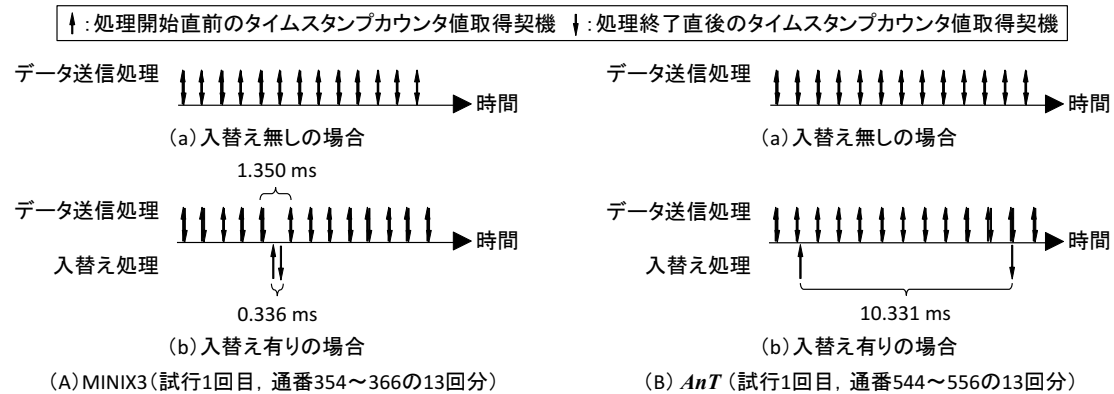


図 2 入替え処理とデータ送信処理の関係

### 3.2 入替え時間

入替え時間とデータ送信時間を表 1 に示す。なお、入替え時間は、1 回 × 試行 5 回 = 5 回、データ送信時間は、1000 回 × 試行 5 回 = 5000 回の測定結果について示している。入替え時間について、表 1 より、以下のことがわかる。

(1) MINIX3 において、データ送信有りの場合の入替え時間の最小値は、データ送信無しの場合の最小値に等しく 0.311 ms である。この時、データ送信処理の影響を受けない契機において入替え処理が行われていることがわかる。

(2) MINIX3 において、データ送信無しの場合と有りの場合の入替え時間の最大値の差は、0.020 ms と短い。これは、入替え時間に対して、データ送信の依頼に対する処理や結果返却に要する時間が短く、これらの処理に入替え処理が待たされる時間が短いためであると推察できる。

(3) *AnT* において、データ送信無しの場合の入替え時間の最大値 0.579 ms に比べ、有りの場合の入替え時間の最大値は、10.331 ms と長い。これは、新 NIC ドライバプロセスの生成時に発生した 10 回のディスク I/O 要求の度に AP プロセスにディスパッチされ、AP プロセスのタイムスライス時間 (1 ms) 終了まで入替え処理が 10 回待たされるためである。

以上より、データ送信有りの場合の入替え時間の最大値について、MINIX3 は、*AnT* より 9.995 ms (= 10.331 ms - 0.336 ms) 短い。ただし、*AnT* の入替え処理に関するプロセスの優先度を上げることで、この差を短縮できると推察できる。

### 3.3 データ送信時間

入替え処理開始契機近傍のデータ送信処理に注目し、入替え処理とデータ送信処理が相互の処理に及ぼす影

響について考察する。

入替え処理とデータ送信処理の関係を図 2 に示す。図 2 より、以下のことがわかる。

(1) 図 2 (A) より、MINIX3 のデータ送信処理の開始間隔は、入替え処理によって最大 1.350 ms (= AP プロセスによる PU 処理時間 1 ms + データ送信時間 0.012 ms + 入替え時間 0.336 ms) となった。これは、入替え処理が行われているため、AP プロセスに CPU が割り当てられず、PU 処理の終了時刻が延びたためである。

(2) 図 2 (B) より、*AnT* のデータ送信処理の 1 回あたりの開始間隔は、入替え処理によりわずかに (平均 0.0794 ms) 増加し、10 回の送信間隔で増加している。この理由は、3.2 節 (3) と同様である。

以上より、入替え有りの場合のデータ送信処理の開始間隔について、*AnT* は、MINIX3 より 0.458 ms (= (0.0794 ms × 10 回) - 0.336 ms) 影響が大きい。

## 4. おわりに

MINIX3 と *AnT* について、NIC ドライバプロセスの入替え時間とデータ送信時間を測定し、入替え処理がデータ送信処理に与える影響について比較評価した。

入替え時間の最大値について、データ送信有りの場合、*AnT* の方が MINIX3 より 9.995 ms 長いことを示した。また、*AnT* の方がデータ送信処理に与える影響が 0.458 ms 大きいことを示した。

## 参考文献

- [1] Andrew S. Tanenbaum, Jorrit N. Herder, and Herbert Bos, "Can we make operating systems reliable and secure?," IEEE Computer Magazine, Vol.39, No.5, pp.44-51 (2006).
- [2] 藤原 康行, 岡本 幸大, 田端 利宏, 乃村 能成, 谷口 秀夫, "*AnT* における OS サーバ入れ替え機能," マルチメディア通信と分散処理ワークショップ論文集, Vol.2008, No.14, pp.201-206 (2008).