

ネットワーク I / O 帯域の隔離

Xen Summit Tokyo 2008

宝曼 西門 <simon@valinux.co.jp>
稲越 宏弥 <inakoshi.hiroya@jp.fujitsu.com>

2008年11月22日~21日

本研究・開発の一部は、経済産業省の委託を受けた
技術研究組合 超先端電子技術開発機構(ASET)のセ
キュア・プラットフォームプロジェクトの成果です

目次

- 第1節：帯域の隔離
- 第2節：パケットの識別
- 第3節：パケットスケジューラ

第1節

帯域の隔離

目的

公平性

- 全部のドメインが公平にネットワークの資源を使えるように
 - 管理者のポリシー設定に応じて
- ウイルスがあるドメインから防衛します
- 悪意あるドメインから防衛します

ネットワークの資源

- NIC帯域
- Dom0 CPU
- Dom0 カーネル内メモリ

ネットワークの資源

- NIC帯域
 - domU送受信速度
- Dom0 CPU
 - domU送受信速度
- Dom0 カーネル内メモリ
 - Dom0 カーネルのバケット待ち時間

パケットスケジューラ

- 送受信したdomUによって、パケットを優先度付けします
- キューは満杯時にパケットを破棄します

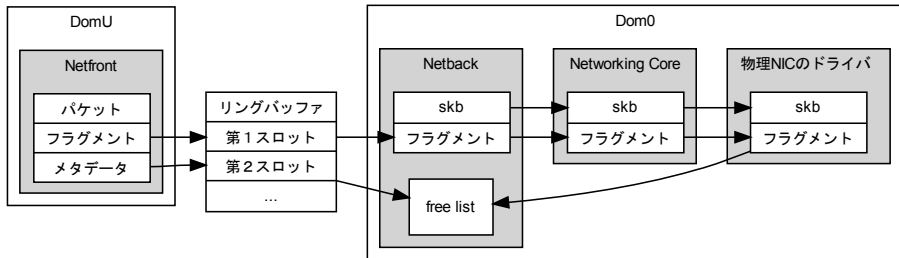
パケットスケジューラ

- 送受信したdomUによって、パケットを優先度付けします
 - NIC帯域
 - Dom0 CPU
- キューは満杯時にパケットを破棄します
 - Dom0カーネル内メモリ

Netback/Netfront フロー制御

Netfrontから物理NICの間でルーティングが行なわれるので、パケットスケジューラはネットワーク資源を制御できます

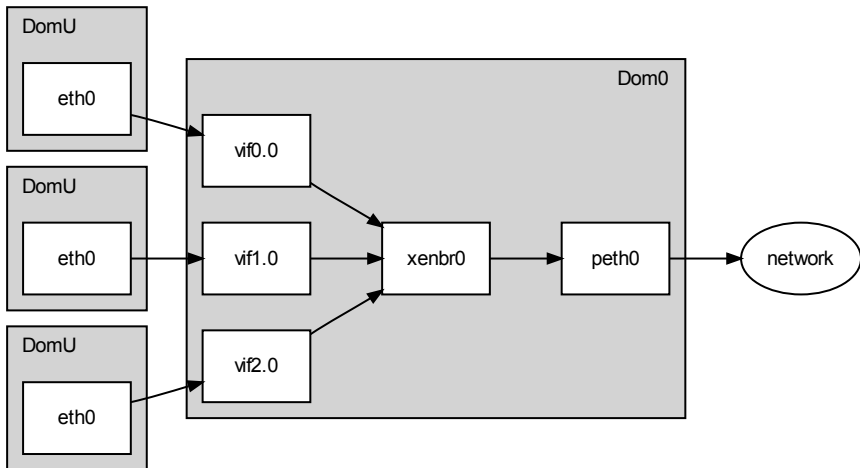
- dom0のCPU
- dom0のカーネル内のメモリ



第2節

パケットの識別

DomU送信: パケットを識別する



- どのインターフェースからxenbr0に転送されたか分かったら送信したdomUが分かります

DomU送信: iptablesのルール

domUが使用するインターフェースによって, fwmarkを付与します

```
iptables -t mangle -A FORWARD -m physdev \  
    --physdev-in vif2.0 -j MARK --set-mark 100  
iptables -t mangle -A FORWARD -m physdev \  
    --physdev-in vif3.0 -j MARK --set-mark 110  
iptables -t mangle -A FORWARD -m physdev \  
    --physdev-in vif5.0 -j MARK --set-mark 120
```

第3節

パケットスケジューラ

パケットスケジューラ

- フィルタリング
 - クラスに対して設定します
- 優先度付け
 - クラス識別に基づきます
 - パケットを遅延させることができます
- キューイング
 - 優先付け時とフィルタリング後、送信のために
- ドロップ
 - キューが満杯の時

Netback/Netfrontのフロー制御

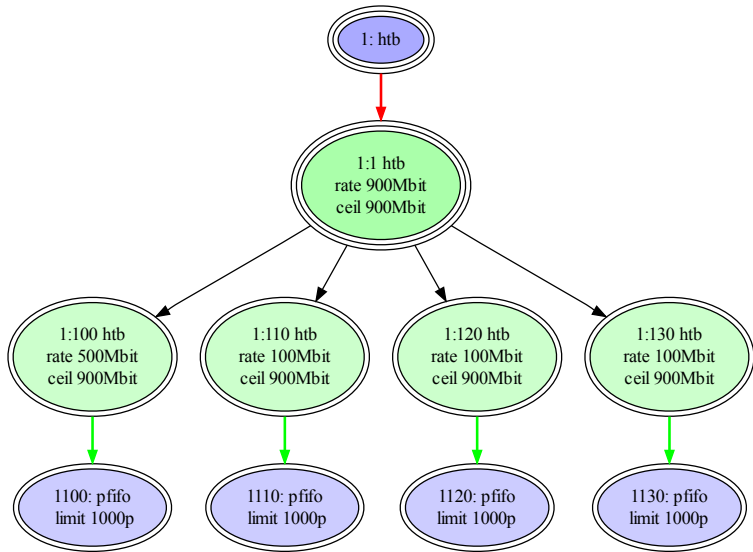
DomU送信

$$p \leq n$$

where: p : vifN.Mから送信したdom0にキューに入れたパケット
 n : netbackのリングバッファの-slot数
(デフォルト= 256)

- dom0にパケットを遅延するので十分です
- dom0内でパケットを破棄するのは悪影響があります
 - dom0内でパケットを遅延させるにしたがって、domU送信は
遅くなります

DomU送信: パケット・スケジューラ の階層図



DomU送信: HTBのルール: Leafクラス

Leafクラス

- ドメインを1個ずつ
- デフォルト

```
tc class add dev peth0 parent 1:1 classid 1:100 htb \  
    rate 500Mbit ceil 900Mbit  
tc class add dev peth0 parent 1:1 classid 1:110 htb \  
    rate 100Mbit ceil 900Mbit  
tc class add dev peth0 parent 1:1 classid 1:120 htb \  
    rate 100Mbit ceil 900Mbit  
tc class add dev peth0 parent 1:1 classid 1:130 htb \  
    rate 100Mbit ceil 900Mbit
```

DomU送信: フィルタ

iptablesによって設定されるfwmarkによりフィルタされます

- handle N はfwmark のキーです
- flowid X:Y flowid X:Yはパケット繋ぐキューです

```
tc filter add dev peth0 protocol ip parent 1: \  
    handle 100 flowid 1:100  
tc filter add dev peth0 protocol ip parent 1: \  
    handle 110 flowid 1:110  
tc filter add dev peth0 protocol ip parent 1: \  
    handle 120 flowid 1:120
```

質問

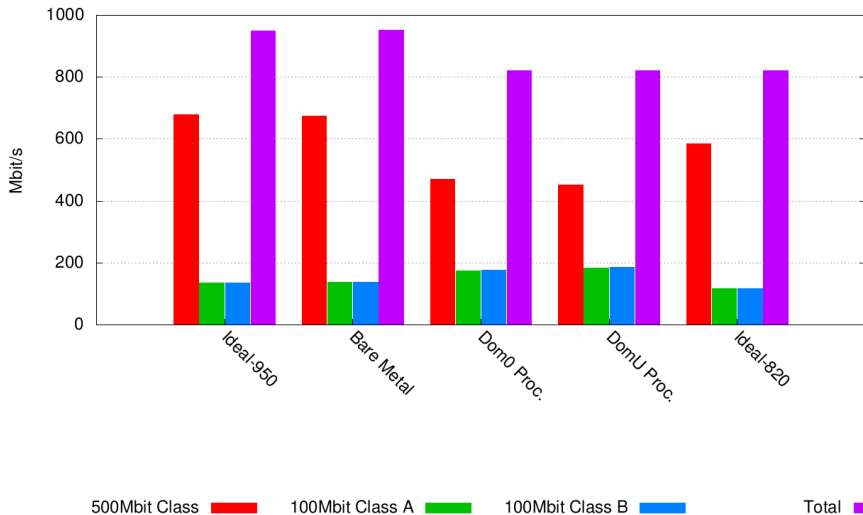
質問がありますか

第4分

延長

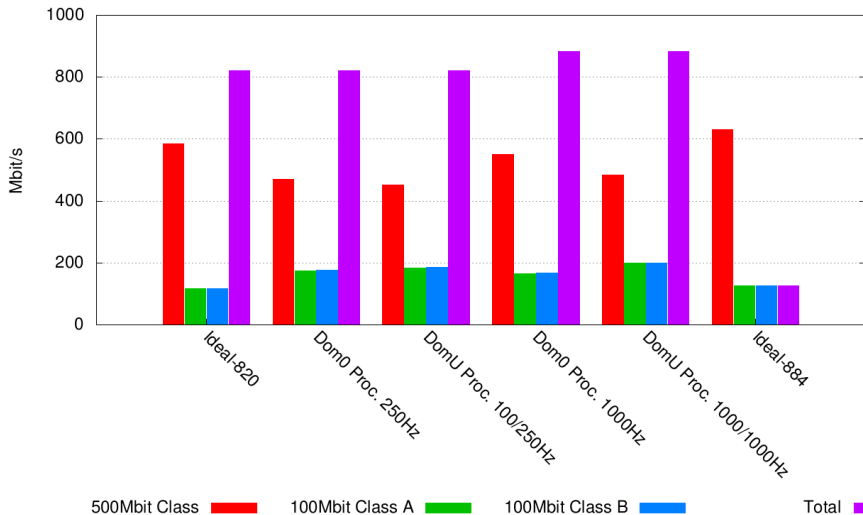
HTBのパフォーマンス

Ubuntu Hardy 2.6.24-21-generic / Xen 3.2 + 2.6.24-21-generic

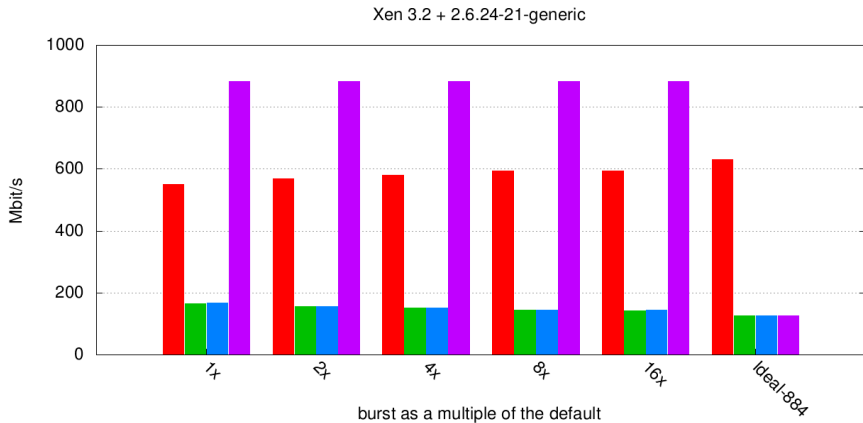


HZのチューニング

Xen 3.2 + 2.6.24-21-generic



バーストのチューニング: Dom0から出るパケット



500Mbit Class



100Mbit Class A



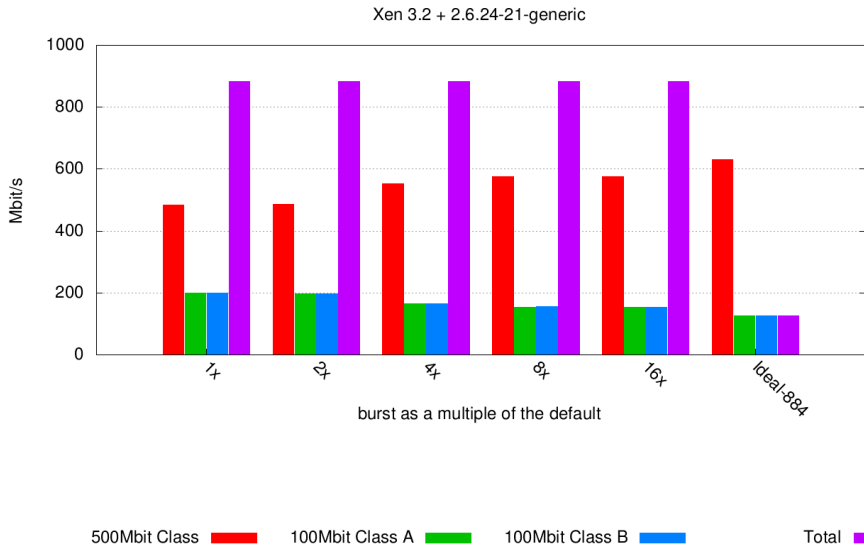
100Mbit Class B



Total



バーストのチューニング: DomUから出るパケット



終了