

Virtual Interface ArchitectureのInternet 拡張方式

1999.06.09

小林 伸治、陣崎 明

新情報並列分散システム富士通研究室

発表内容

- 並列分散システム高速化の課題
 - Internet環境(IP)の利用
- VIAについて
 - Internet環境におけるVIA利用の課題
- VIAのInternet拡張、Comet-VIA
 - 方式説明
 - 実装の説明
 - 評価結果
- まとめ

並列分散システムの高速度化

- プロセッサ数の大規模化
 - LANによる結合
 - 高速広域ネットワークによる結合

広域ネットワークによる結合

- 専用回線で結合
- 既存の広域ネットワーク利用
 - Internet環境の利用(IPの採用)が不可欠



並列分散システムの高速度化

- プロセッサ数の大規模化
 - LANによる結合
 - 高速広域ネットワークによる結合
- I/Oの大規模化
 - Storage Area Network

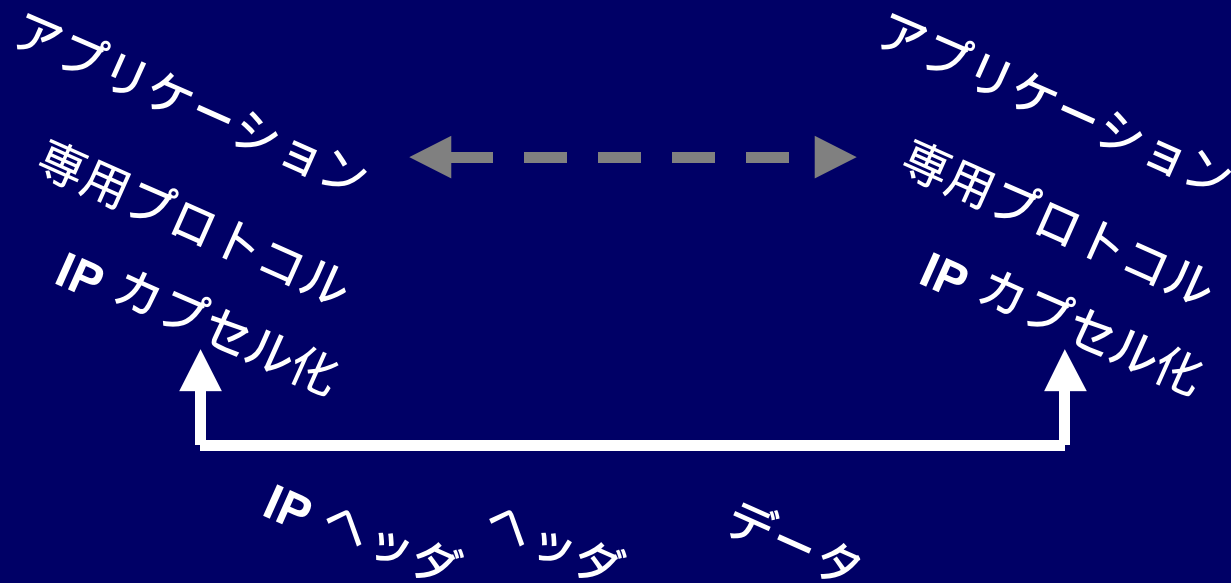


- IP (Internet Protocol) の利用

並列分散システムでのIPの利用

- API
 - socket を利用
 - アプリケーションの書き換え
 - 専用プロトコル over IP
- 重いプロトコル処理
 - プロトコル処理のハードウェア化
 - プロトコルの変更に対応できるか

専用プロトコル over IP

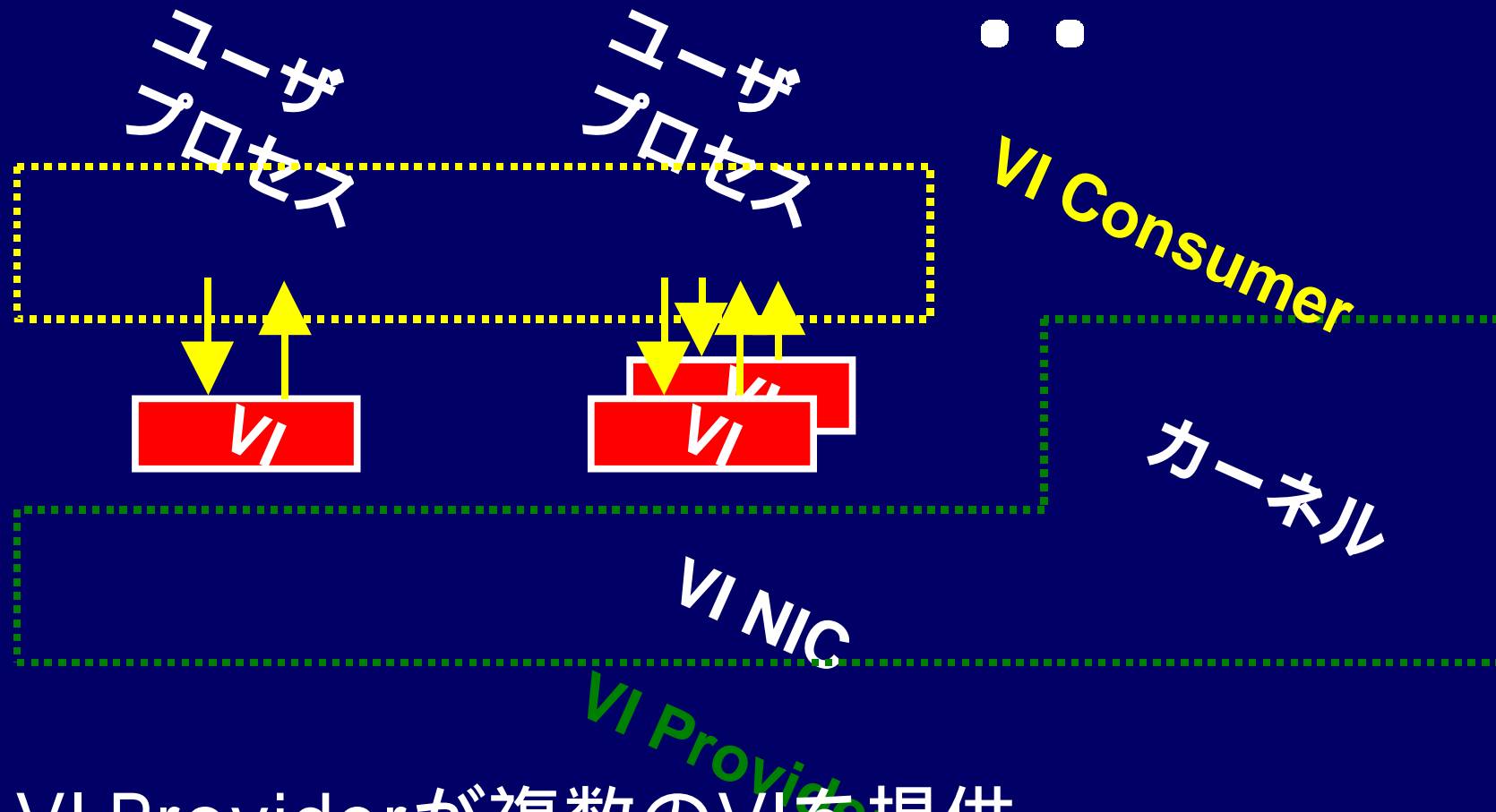


- 専用プロトコルのパケットをIPで搬送
- APIは変更不要

VIA (Virtual Interface Architecture)

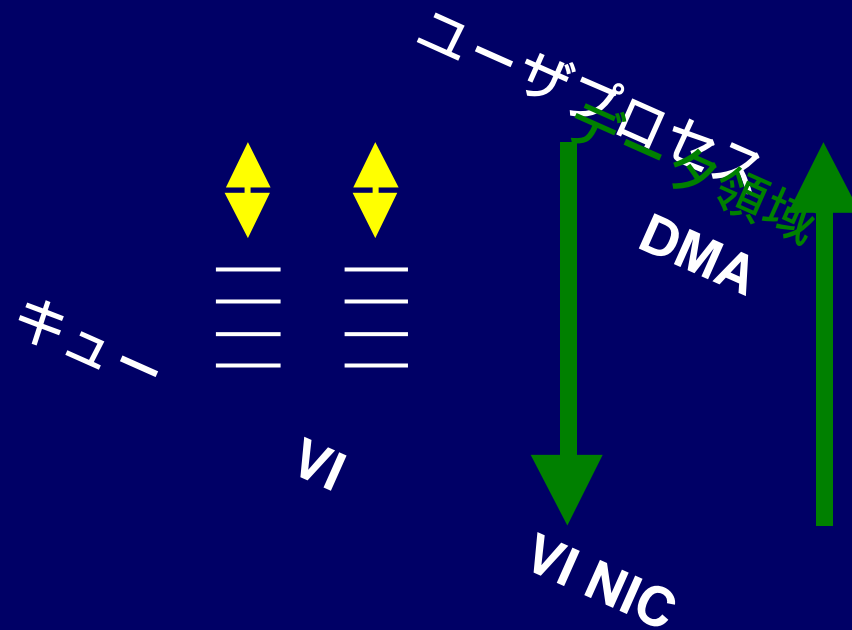
- ユーザレベルネットワークインタフェース
- 標準化の動き
 - Compaq、Intel、Microsoft
- VIAの仕様
 - ネットワークインタフェースの仮想化
 - VI (Virtual Interface)
 - VI間の通信プロトコル
 - コネクション指向

VIA(2)




- VI Providerが複数のVIを提供
- ユーザプロセス内のVI ConsumerがVIを利用

VIA(3)

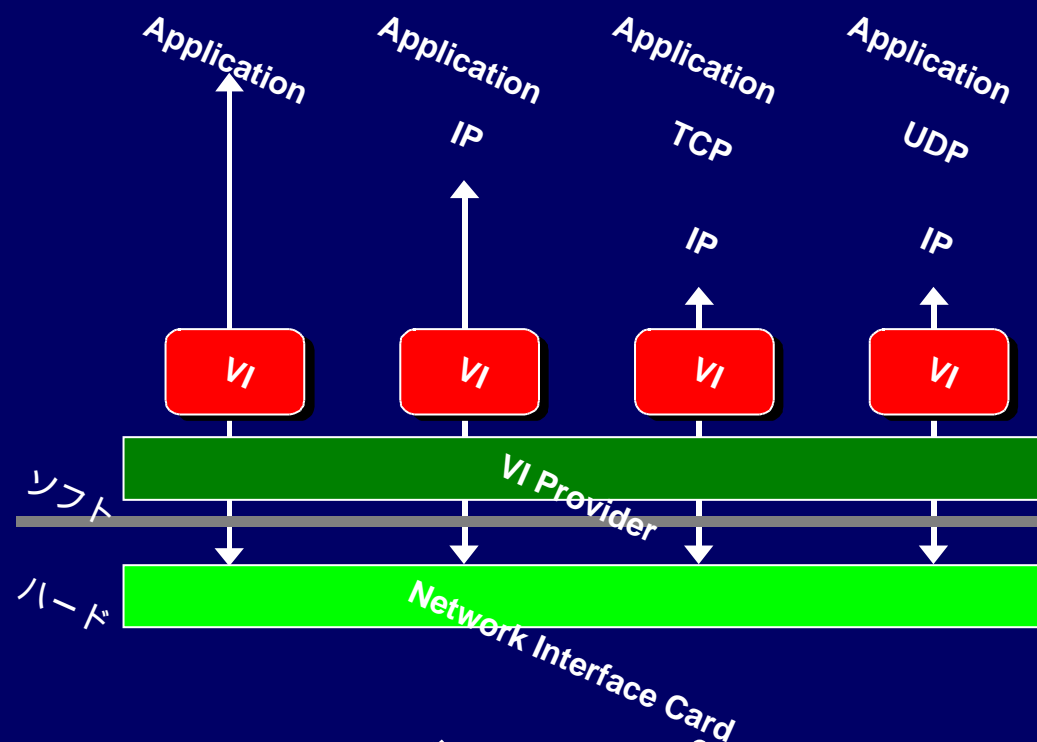


- ユーザプロセスはデータ領域を登録
- VI毎のディスクリプタキューでコマンド発行
- VI NICがデータ領域に直接アクセス(DMA)

VIAの利点

- 高速性
 - カーネル介在無し
 - リソース競合の制御
 - 各ユーザプロセスがVIを専有できる
 - ディスクリプタキューはユーザプロセス中に存在
- 
- パケット廃棄が起こらないようにできる
 - 信頼性の確保

Internet環境におけるVIA利用



- 1種類のVIでさまざまなプロトコルに対応
 - ユーザ空間でのプロトコル処理
- 高機能ネットワークアダプタを活かせない

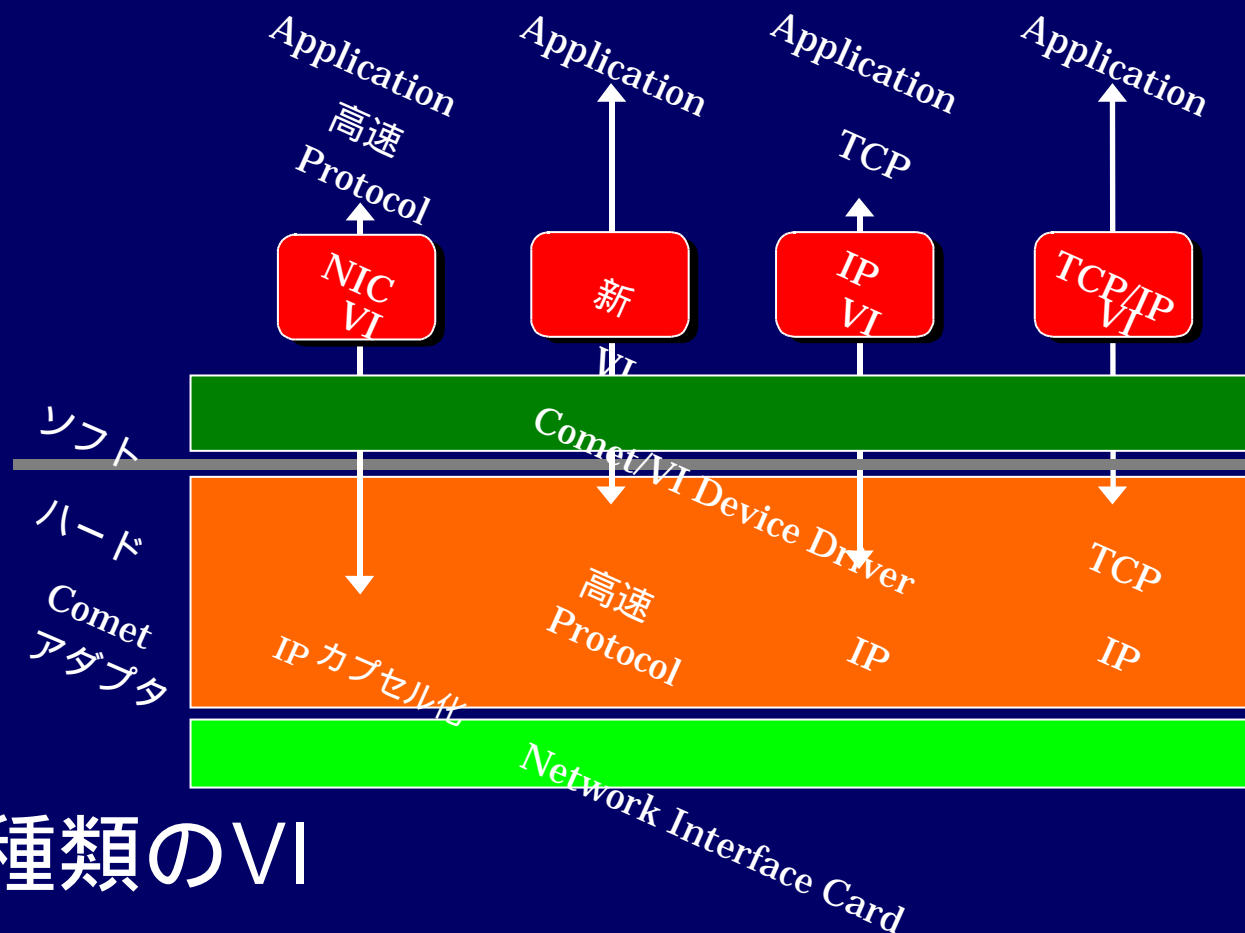
Comet-VIA

- VIAの問題点
 - さまざまなプロトコルに対応しづらい
 - 高機能ネットワークアダプタを活かせない



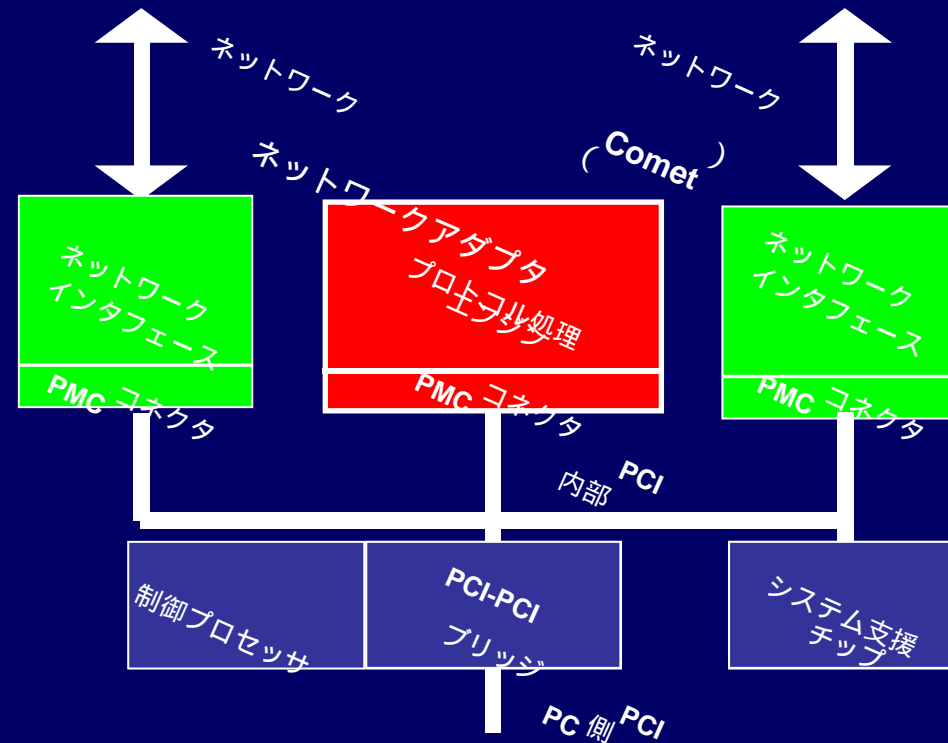
- Comet-VIA方式
 - 複数種類のVIを提供
 - さまざまなプロトコルに対応
 - ネットワークアダプタ内でプロトコル処理
 - ハードウェアによる高速処理

Comet-VIAのモデル



- 複数種類のVI
 - 各アプリケーションが最適なVIを利用
- ネットワークアダプタ内でプロトコル処理

Cometネットワークアダプタ

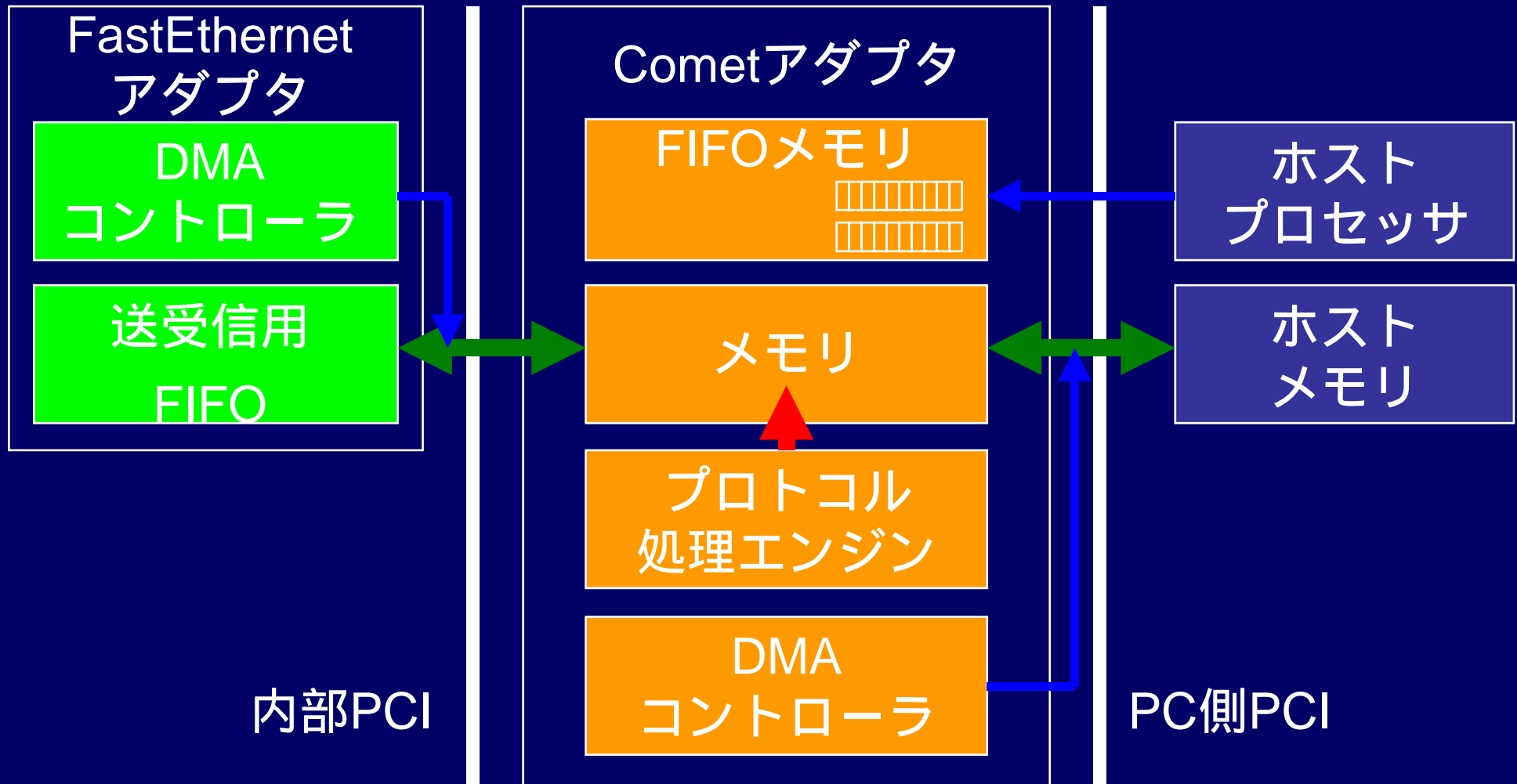


- プロトコル処理エンジンを持つ
- 現試作は汎用プロセッサでエミュレート

Comet-VIAの実装(1)

- 目的
 - 「専用プロトコル over IP」の実現性確認
 - 性能評価
- 実装した機能
 - ライブラリ
 - 専用プロトコル、Comet-FM
 - VI Provider Library
 - ファームウェア
 - Comet-FMパケットのIP化
 - Ethernet ドライバ

Comet-VIAの実装(2)



Comet-FM

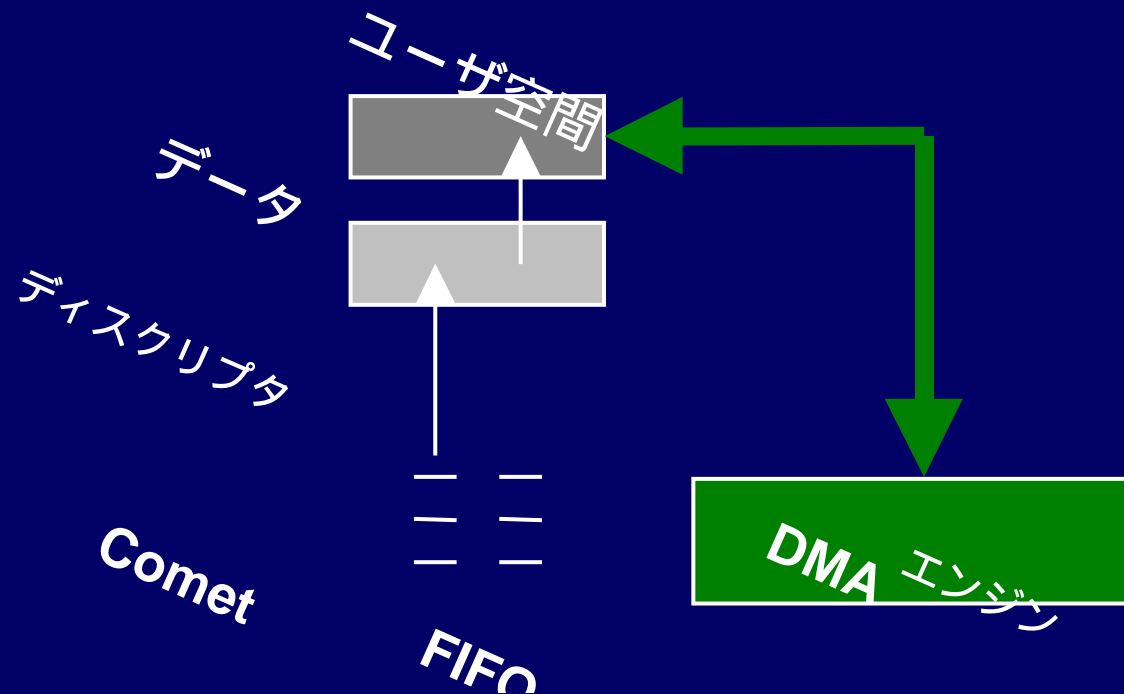
- Illinois大学 Fast Messages 1.1 ベース
 - 送信側が受信ハンドラを指定
- ノンブロック送信プリミティブを追加

プリミティブ	機能
FM_register()	送信バッファ領域登録
FM_deregister()	送信バッファ領域解放
FM_send()	送信 (ノンブロック)
FM_send_nw()	送信完了待ち
FM_wait()	受信ハンドラ呼び出し
FM_extract()	

VI Provider ライブラリ

- Comet-FM over IPに必要な機能を実装
- CometネットワークアダプタのFIFOメモリをDoorbellとして利用
- CometネットワークアダプタはMMUを持たない
 - ライブラリ内で仮想アドレスから物理アドレスに変換
 - 変換キャッシュ

FIFOを利用したDoorbell

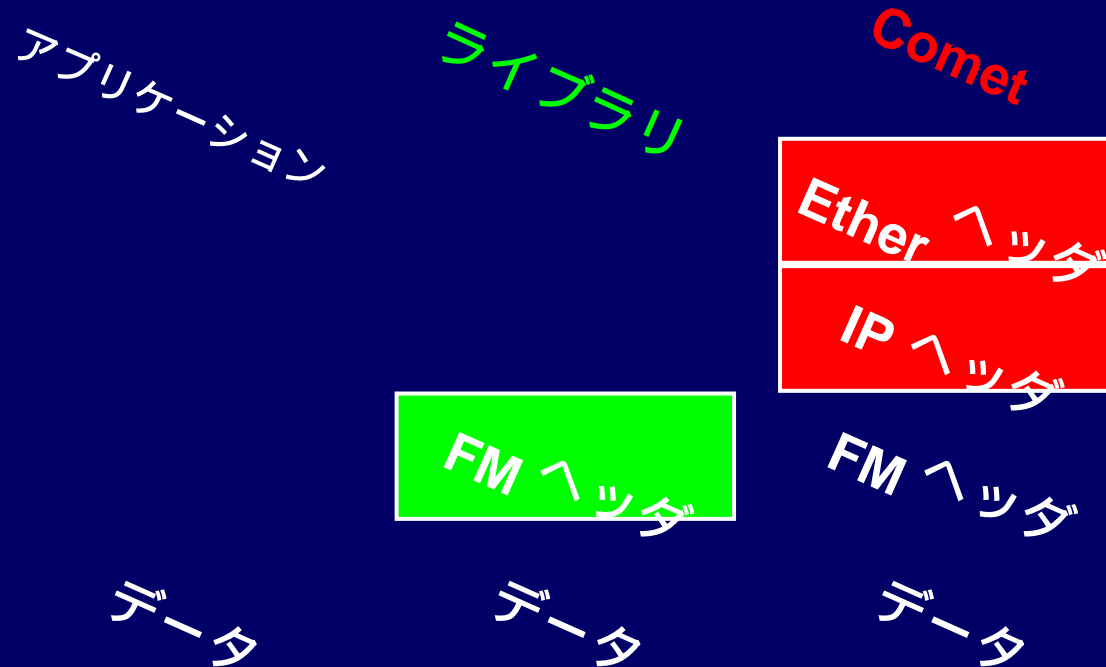


- ディスクリプタの物理アドレスをFIFOメモリで指示
- ディスクリプタはデータの物理アドレスを指示

ファームウェア

- FIFOメモリを監視しホストとのデータ送受信
- 送信時
 - Comet-FMヘッダからIPヘッダ、Ethernetヘッダを生成
 - ノード番号とIPアドレスの対応表
 - ハードウェア化可能
- 受信時
 - IPパケットからComet-FMパケットを抽出
 - ノード番号に対応したVIに渡す
 - ノード番号とVIの対応表
 - ハードウェア化可能

Comet-FM over IPの処理区分



- アプリケーションは「Comet-FM over IP」を意識しない

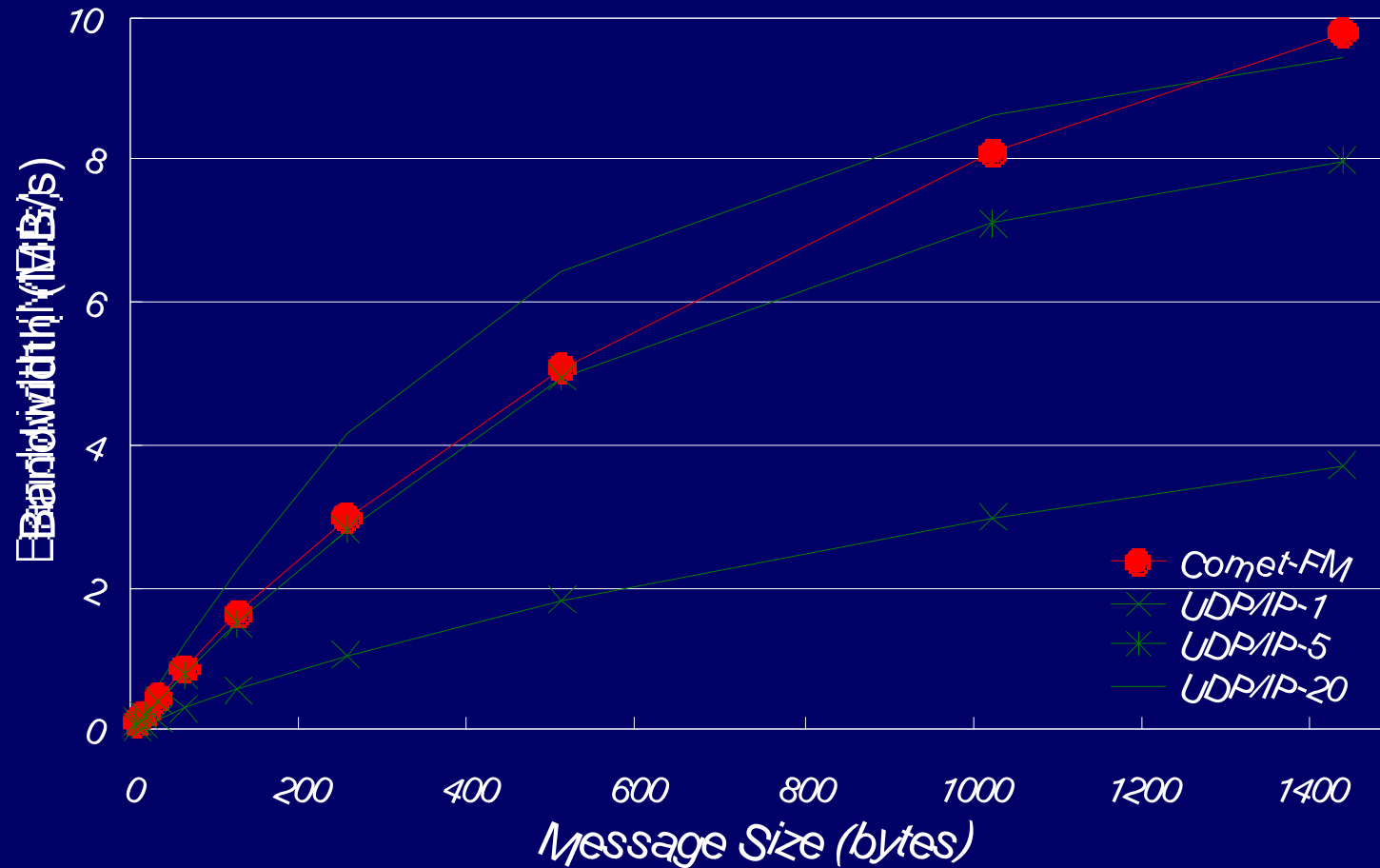
信頼性の確保

- ノード内
 - UDP/IPはノード内でも廃棄が起こる
 - 確認応答必須
 - Comet-VIAでは各プロセスがVIを専有
 - 資源競合が発生しない
 - パケット廃棄無し
- ノード間
 - LAN環境ではほとんど廃棄されない
 - 広域では資源予約で対処可能

性能評価

- FastEthernetで2ノードを結合
 - PentiumPro 200MHz、BSD/OS 3.1
 - UltraSPARC-II 250MHz、Solaris 2.6
- 連続送信時のバンド幅を測定
 - UDP/IPは1、5、20メッセージ毎に確認応答
 - Comet-FMは確認応答不要

測定結果(バンド幅)



- メッセージサイズ大: 高バンド幅
- UDP/IPは確認応答の頻度で変化

まとめ

- VIAをInternet環境で利用する際の問題点
 - さまざまなプロトコルに対応しづらい
 - 高機能ネットワークアダプタを活かせない
- Comet-VIA方式を提案
 - 複数の種類のVIを提供する
 - ネットワークアダプタ内でプロトコル処理
- 「専用プロトコル over IP」を実装
 - 高バンド幅を実現できることを確認
 - 並列分散システムのIPによる高速結合手段として利用可能