

CometにおけるVIA的インタフェース

98.08.05

小林 伸治、陣崎 明

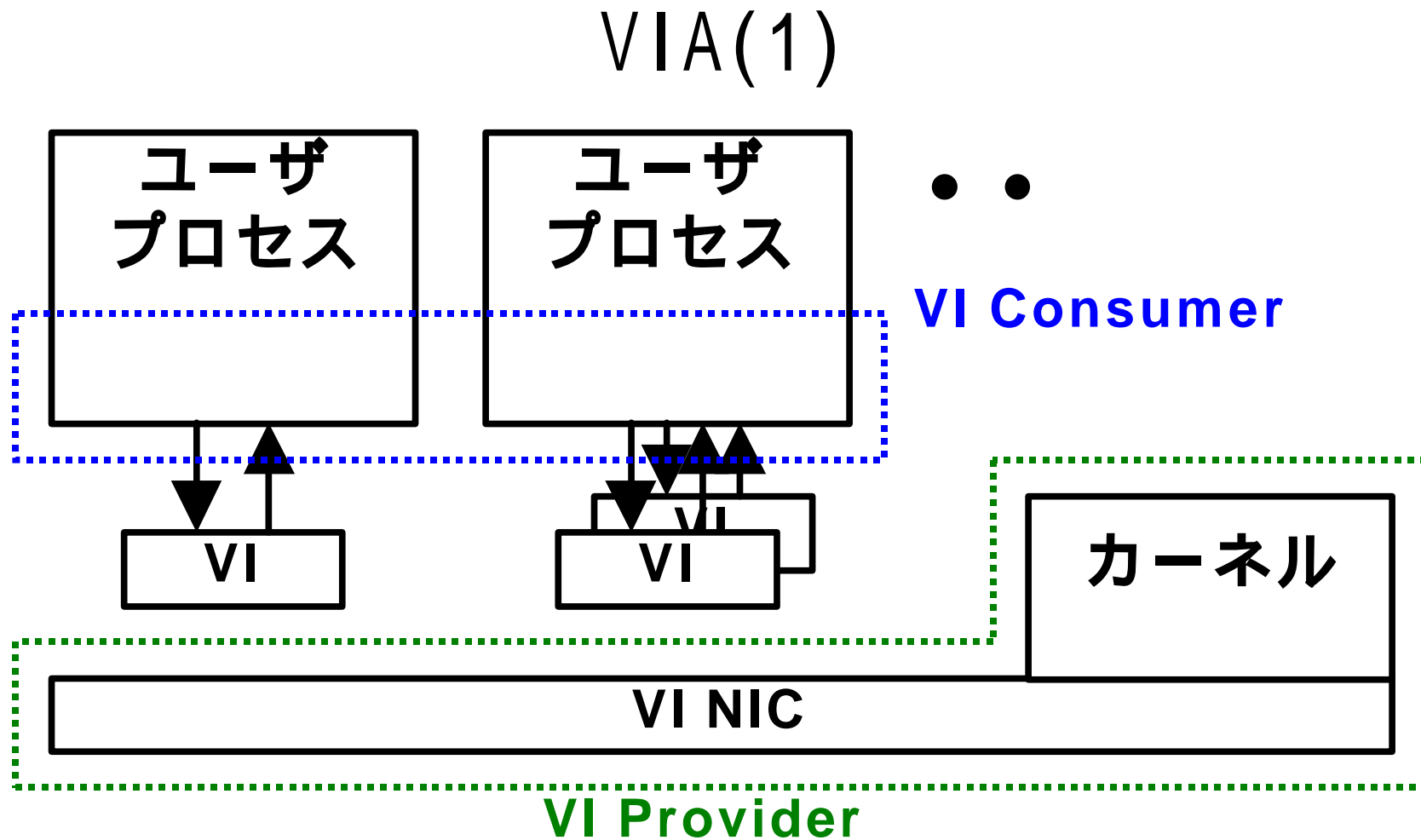
新情報並列分散システム富士通研究室

発表内容

- ネットワークAPIの課題
- VIAについて
- CometにおけるVIAの利用
 - VIA over TCP/IP
 - TCP/IPノードと通信可能なVI
- まとめ

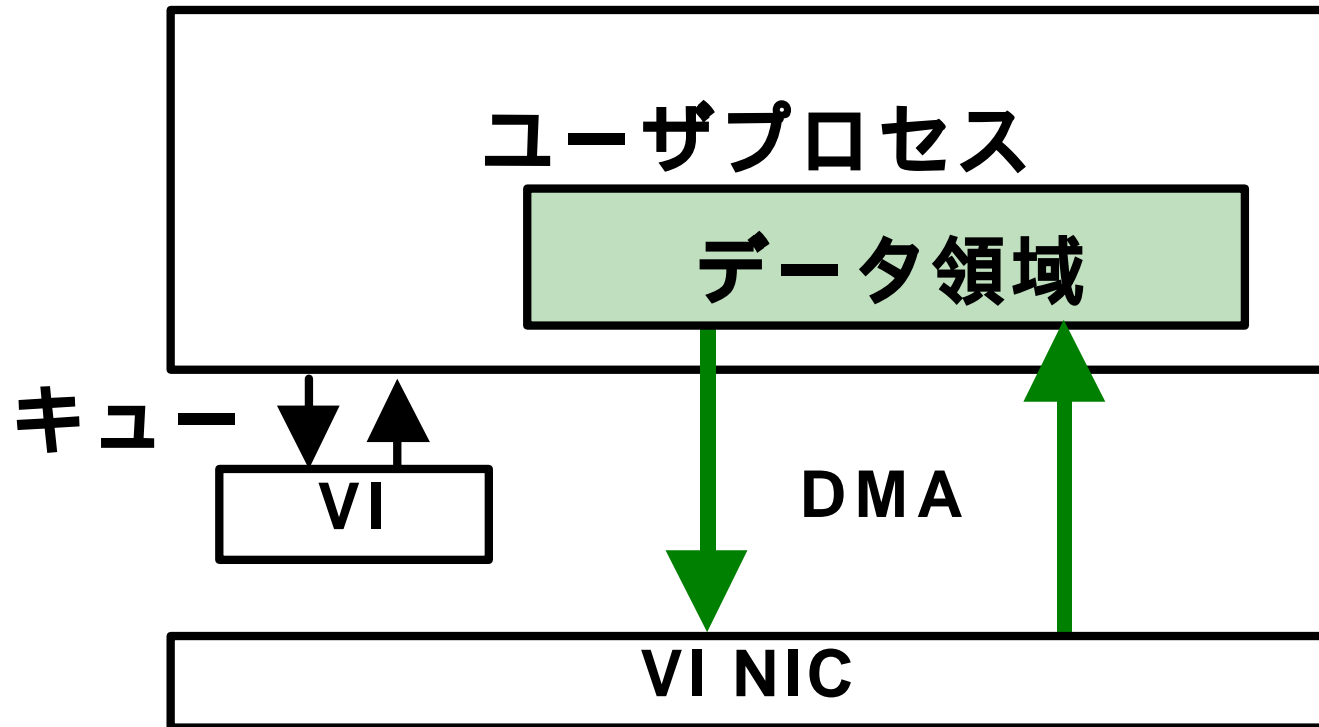
ネットワークAPIの課題

- 要件
 - 高速性
 - 並列計算を実行可能な低遅延性
 - 広域性
 - TCP/IPサポートが不可欠
 - 汎用性
 - 並列計算クラスタとしても高速ネットワークサーバとしても利用
- ユーザプロセスマップ方式は有効か?
 - VIA (Virtual Interface Architecture)
- CometアダプタにVIAは有効か?



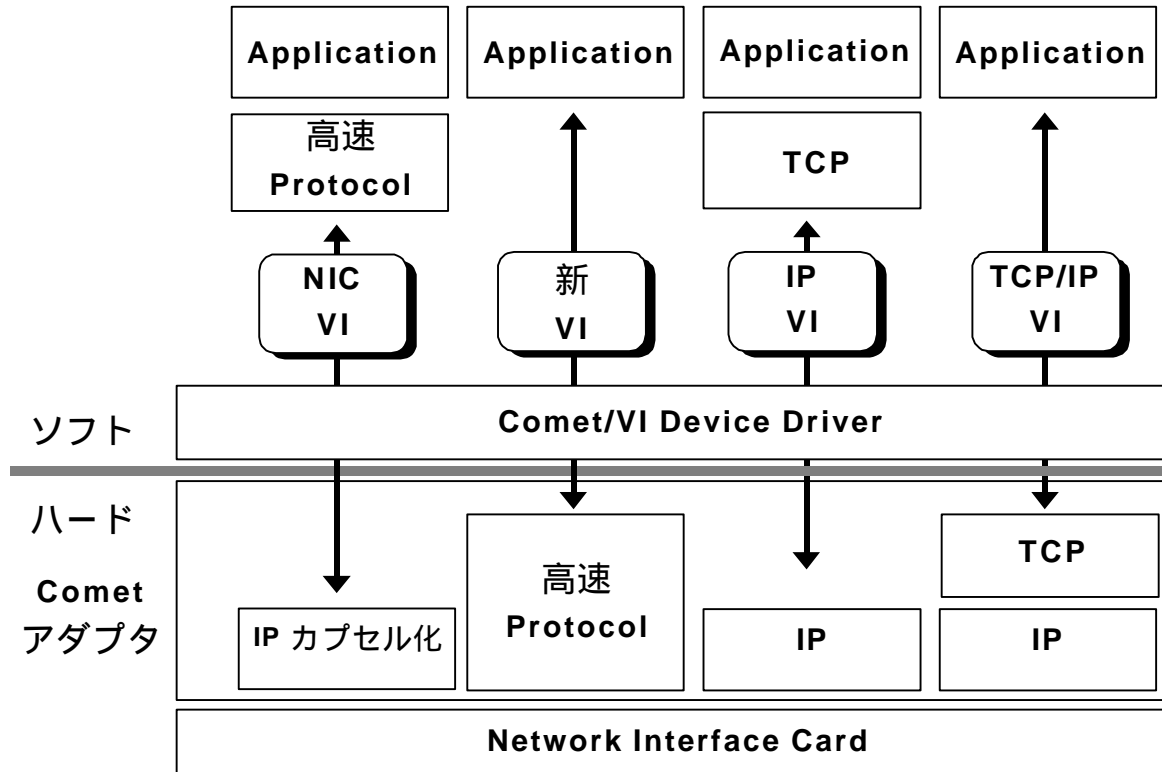
- VI Provider (VI NIC+カーネル) が複数のVIを提供
- ユーザプロセス内のVI ConsumerがVIを利用

VIA(2)



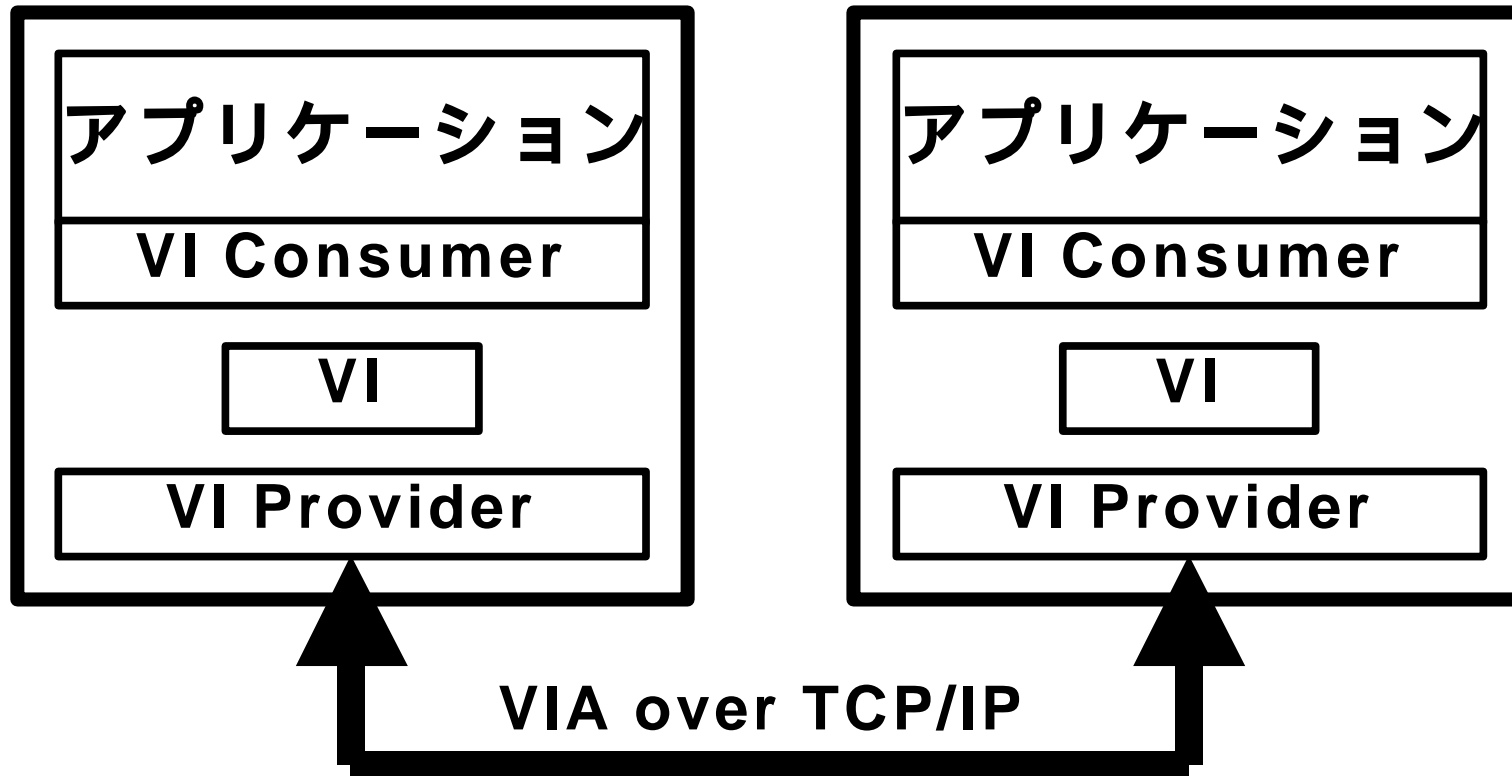
- ユーザプロセスはデータ領域を登録
- 各VI毎に用意されるキューを介してコマンド発行
- VI NICがユーザプロセス中のデータに直接アクセス(DMA)

CometにおけるVIAの利用



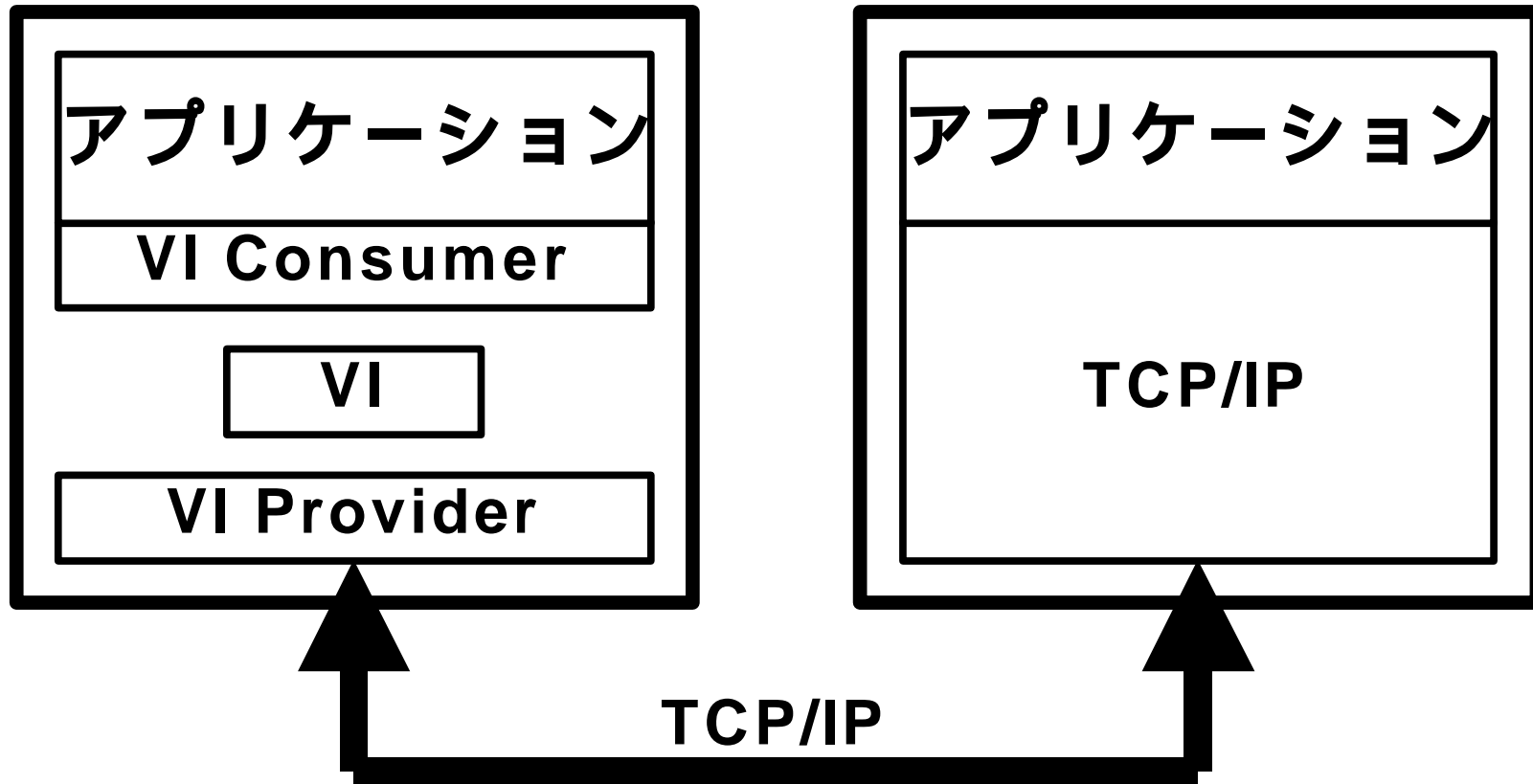
- 方式1 :VIA同士の接続にTCP/IPを利用 (VIA over TCP/IP)
 - 広域分散並列計算
- 方式2 :TCP/IPノードと通信できるVIを提供
 - 高速ネットワークサーバ

方式1のモデル



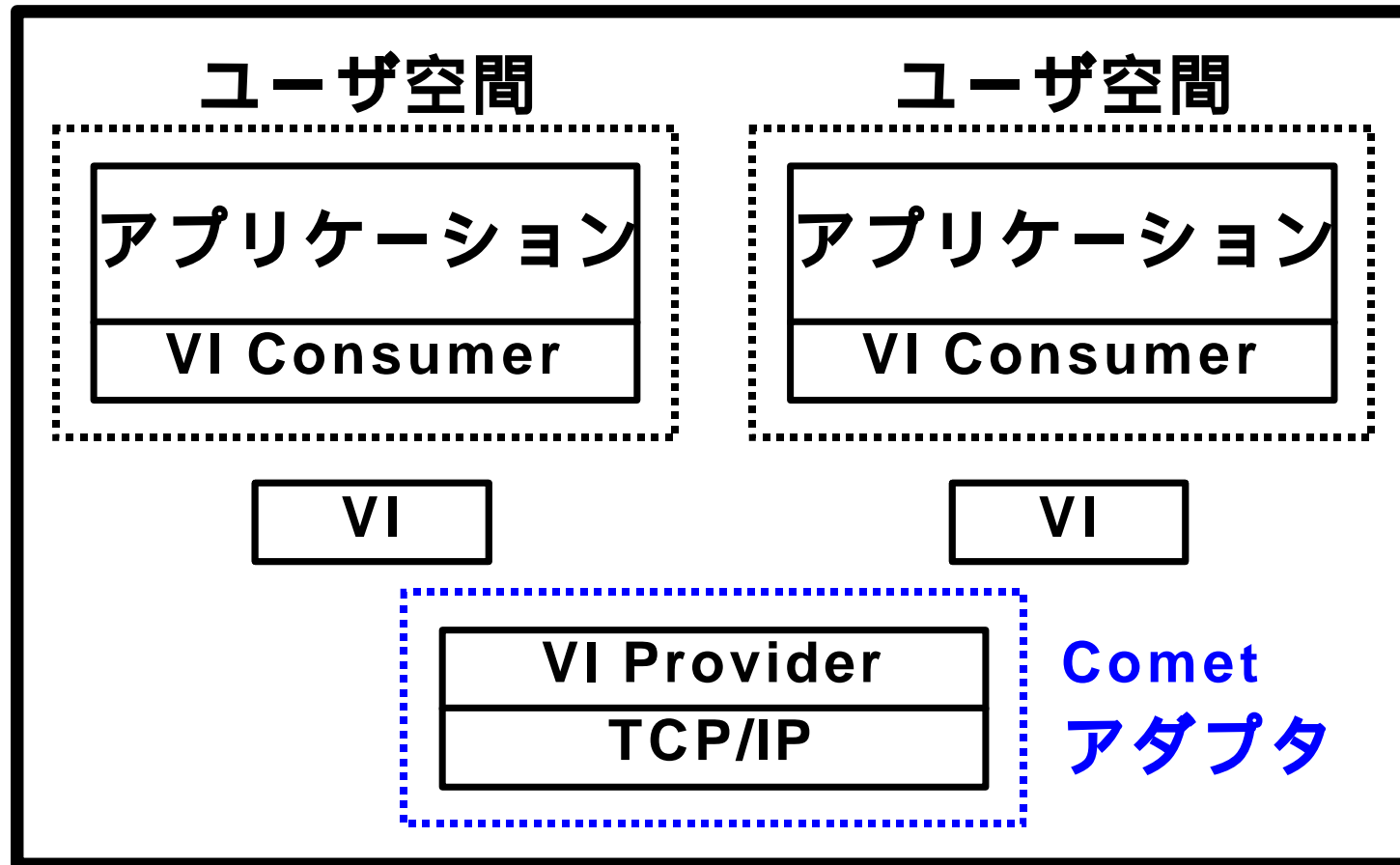
- VIA同士の通信にTCP/IPを利用
- VIAのパケットをIPカプセル化

方式2のモデル



- TCP/IPが使える口をVIAに用意

Cometでの方式1の実現



- Comet アダプタ内でVI間通信をIPパケット化

方式1の検証(1)

- VIAで要求される信頼性

信頼性のレベル	Unreliable	Reliable Delivery	Reliable Reception
データ誤りの検出	要	要	要
重複配送しない保証	要	要	要
データ配送の保証	不要	要	要
データの順序性保証	不要	要	要
送信要求の完了	送信後	送信後	受信後

- TCP/IPが提供する信頼性

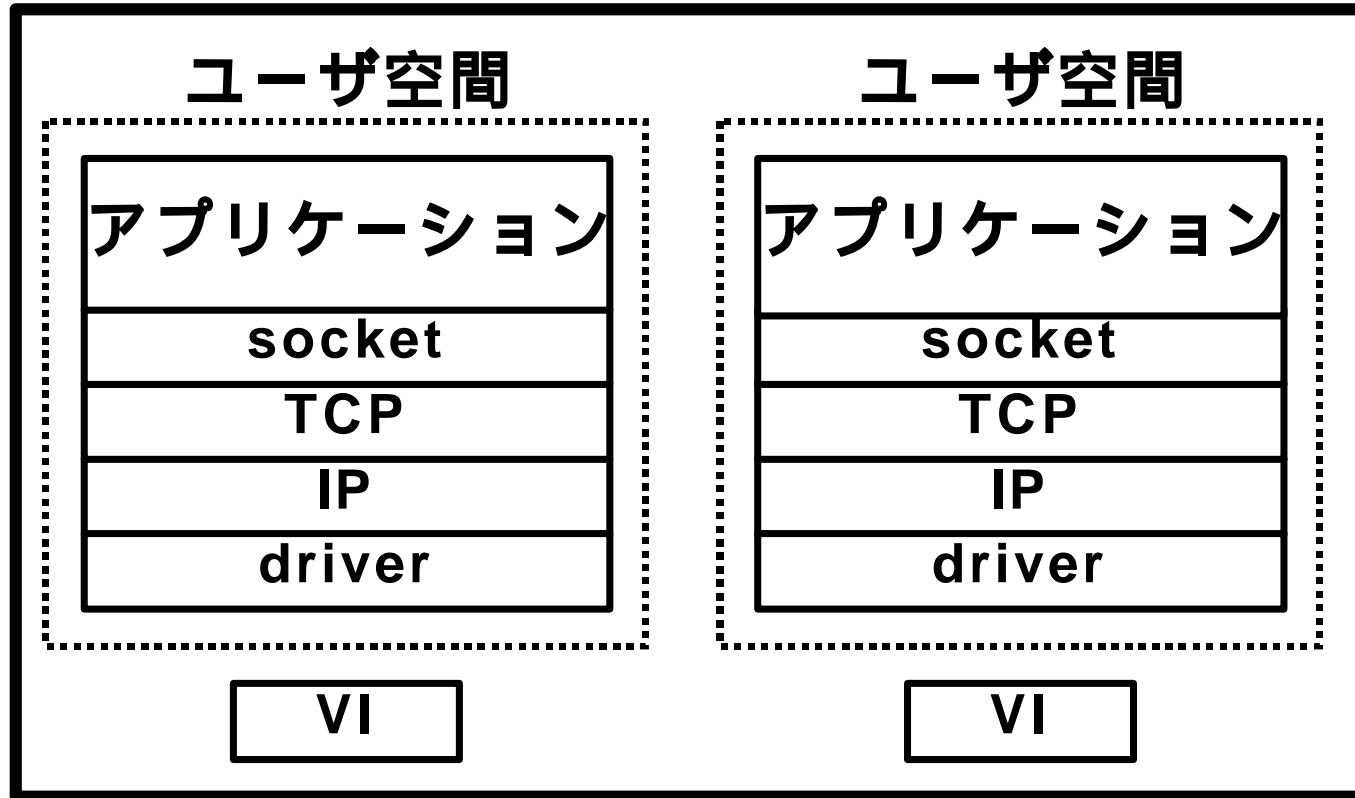
プロトコルの種類	IP	UDP	TCP
データ誤りの検出	無	有	有
重複配送しない保証	無	無	有
データ配送の保証	無	無	有
データの順序性保証	無	無	有

方式1の検証(2)

- アドレス
 - VIAのアドレスは任意長
 - 上位ライブラリ等のアドレス体系やPアドレスをそのまま使用可能
- 遅延
 - IPカプセル化は現状でも40 μ 秒程度 ハード化により5 μ 秒程度
 - 任意のアドレス体系をPアドレスにマッピングする処理も軽い
 - FastMessages実装時の送信遅延予測

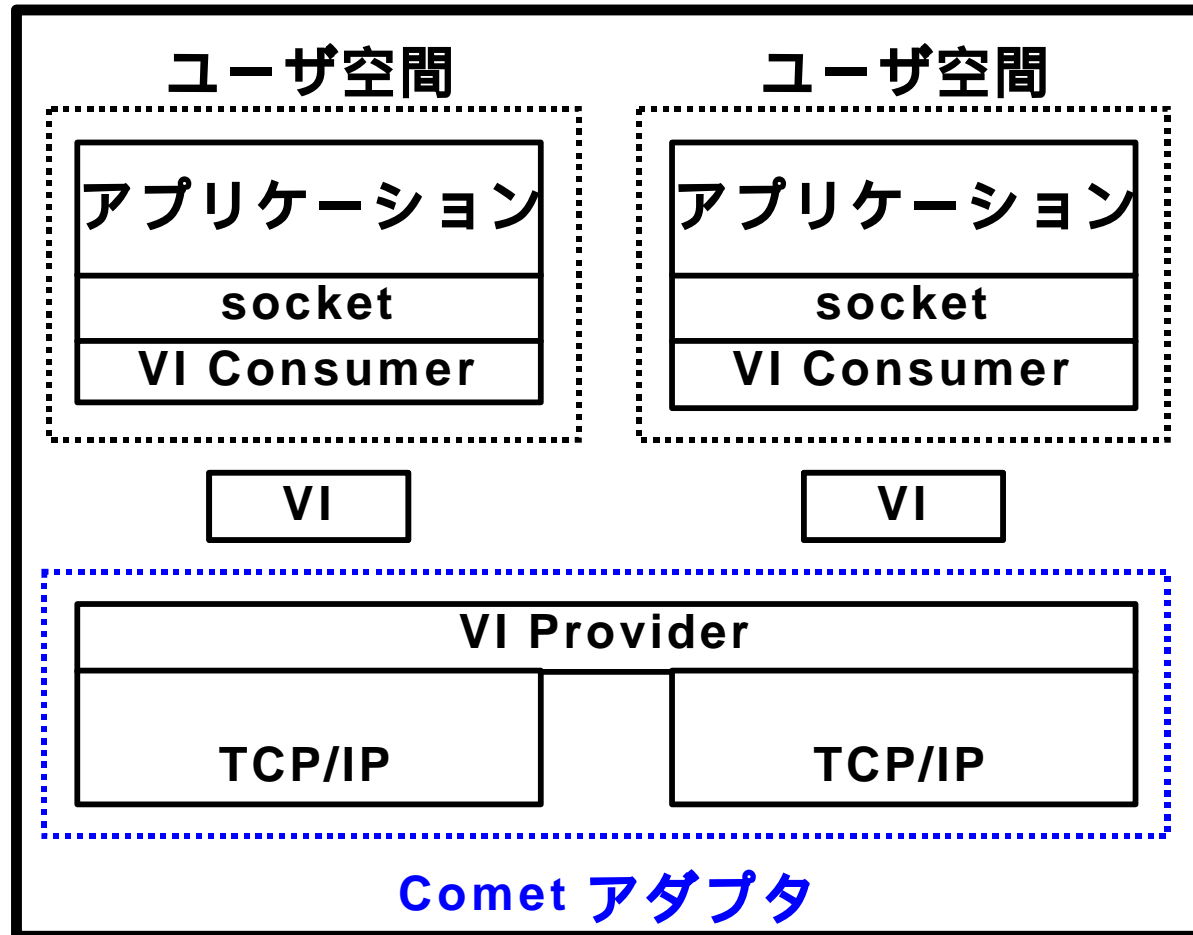
	FastMessages	IPカプセル化	合計遅延
SPARCstation 2 (80MHz) 実測値	20 μ 秒	-	-
PentiumPro 200MHz 予測値	5 μ 秒	5 μ 秒	10 μ 秒

方式2の一実現



- ドライバ層がVIを利用
- TCP/IPプロトコル処理はメインプロセッサが担当

Cometでの方式2の実現



- TCP/IPプロトコル処理はCometアダプタが担当

まとめ

- CometのソフトウェアインタフェースとしてVIAを検討
 - VIAにより様々なプロトコルを統一的に扱うことができる
- CometによるVIA over TCP/IPの特性を検討
 - VIAが求める特性を実現できることを確認
- VIAはCometアダプタのネットワークAPIとして有望
- 今後実装を行う