JDSF　FCT部会　勉強会
Infinibandプロトコルアナライザ

2004年5月12日

株式会社 東陽テクニカ
I/Oアクセス解析ソリューション部
堀部　勝義

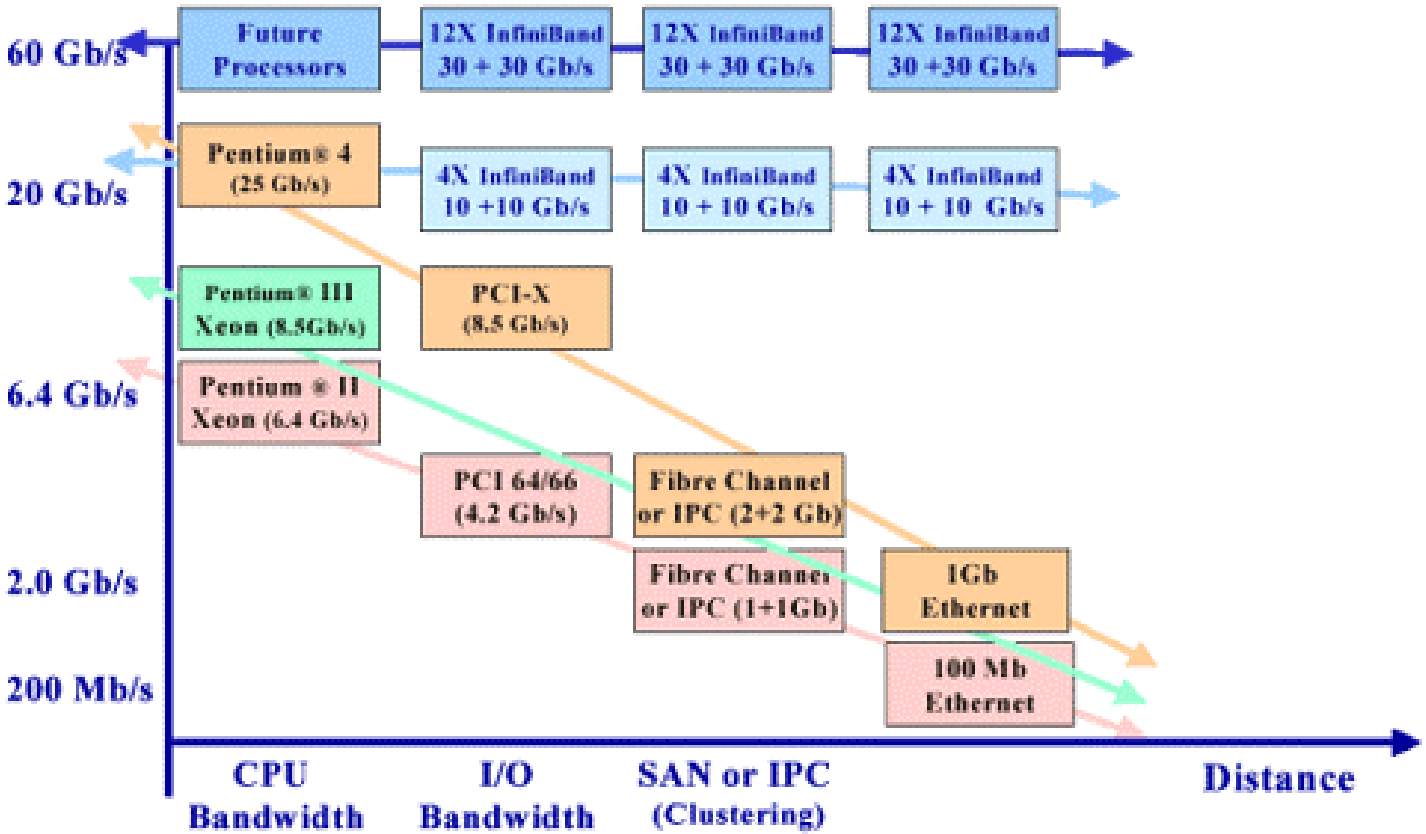
Infinibandプロトコルアナライザ

- 内　容
  - Infinibandの概要
    - プロトコルのロードマップ
    - Infinibandの特徴
    - Infinibandの技術概要
  - Infinibandプロトコルアナライザについて

株式会社 東陽テクニカ

Infinibandの概要

Infinibandロードマップ



株式会社 東陽テクニカ

Infinibandの概要

Infinibandの特徴

- 階層化されたプロトコル
- マルチ・レイヤーの接続性
- パケットベースのコミュニケーション
- マルチキャスト機能
- パケットおよびエンド・ノードのフォルト・トレランス
- サブネット・マネージメント機能
- リンクスピード - 1x, 4x, 12x
- 2.5 ~30 Gbit / sec.の転送レート
- PCB、カッパー、ファイバーのフィジカル・リンク
- リモートDMA をサポート

Infinibandの概要

- **Channel-Based Architecture** – InfiniBandアーキテクチャはチャンネルベースの I/O モデルにより、ファブリックノード間で信頼性の高い接続が可能
- **Message-Passing Structure** – InfiniBandアーキテクチャ・プロトコルは データ転送に効果的なmessage-passingストラクチャを採用
- **Natural Redundancy** - InfiniBandファブリックでは、ノードはlink redundancyをファブリックにアタッチできるので、1つのパスが失敗しても、トラフィックは最終的なエンドポイントに再ルート可能。
- **Quality of Service (QoS)** – リンクレイヤはInfinibandのQoS 特性を可能にします。InfiniBand は15の独立したレベル (VL0-14)とひとつのマネージメントパス (VL15) でデバイス特定の優先順位を構成することが可能. これによりファブリックにおけるI/Oオペレーションに優先権がアサインされクリティカルな通信をすることが可能。
- **Credit-based flow control** –クレジット・ベースのフロー制御管理アプローチにより、IBAネットワーク上の各受信ノードは、パケットロスなしに効率的に転送可能な最大データ量の値を送信デバイスに転送する。クレジット・データは、IBAファブリックに沿って専用のリンク経由で転送。送信デバイスは受信デバイスからプライマリ通信バッファ経由でデータ転送が可能であることを知るまで、パケットは送信しない。
- **CRC Check** – 2つのCRCによりエンド to エンドのインテグリティをチェック; 16-bit variant CRC値は各データ・フィールドに指定されて、各IBAファブリック・ホップで再計算される。32ビットinvariant CRC値は、各IBAホップポイント間で変わらない静的データを保護するように設計されている。
- **Subnet Management** – InfiniBandのネットワークレイヤは サブネット間のパケットルーティングを提供。ルートされた各パケットは、ソースとデスティネーション・ノードのためにグローバル・ルート・ヘッダ (GRH) と128ビットIPv6アドレスがある。またネットワークレイヤは、全てのサブネットに沿って各デバイスのために標準64ビットのユニークなグローバルな識別子を埋めます。これらの一貫した識別値のハンドリングにより、IBAネットワークは、複数の論理的サブネット経由のデータ転送が可能。

株式会社 東陽テクニカ

Infinibandプロトコルアナライザ

内　容

- Infinibandの概要
  - プロトコルのロードマップ
  - Infinibandの特徴
  - Infinibandの技術概要
- Infinibandプロトコルアナライザについて
  - IB Tracer 1X アナライザ
  - IB Tracer 4X アナライザ
  - IB Trainer 4X エキササイザ
  - Infiniband SPEC1.1のサポート
  - Lane Reversalについて
  - Script Verification Engineについて
- 東陽テクニカ　ストレージアプリケーション用プロトコル解析評価ツール製品ライン

株式会社 東陽テクニカ

IBTracer×4プロトコルアナライザ

Infiniband 4xのプロトコル解析

- Infiniband Spec1.1**をサポート**
- InfiniBand 1X(2.5Gbps), 4X(10Gbps)**転送速度のバストラフィックを記録**

株式会社 東陽テクニカ

IBTracer×1プロトコルアナライザ

Infiniband 1xのプロトコル解析

- Infiniband Spec1.0**をサポート**
- InfiniBand 1X(2.5Gbps),**転送速度のバストラフィックを記録**

株式会社 東陽テクニカ

IBTracer　Infinibandプロトコルアナライザ

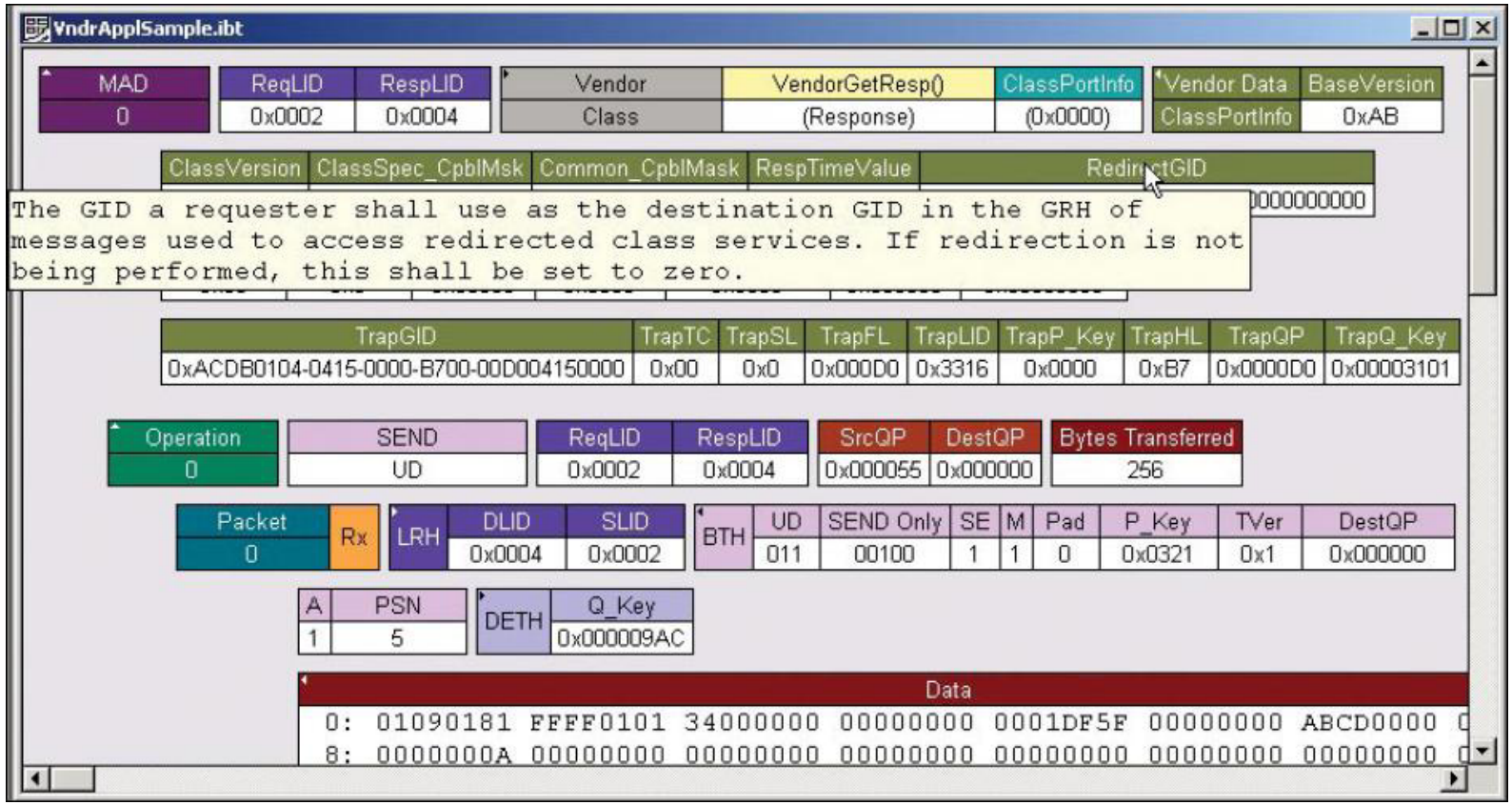
IBTracerの特徴

- ドリル・ダウン表示：MAD（マネージメント・データグラム）、エラー、ペイロード、パケット
- マルチレーンリンク解析：late-to-lane スキューの記録と表示
- Lane-Reversal 対応：レーンの反転をトリガ、記録、表示対応
- SPEC1.1の新しいMAD タイプをサポート
- Script Verification Engine：ユーザ定義可能なスクリプトにより自動解析が可能
- トレーニング・シーケンスを表示
- スキップ・オーダードセット、フラグ・バイオレーション を表示

株式会社 東陽テクニカ

IB Tracer　Infinibandプロトコルアナライザ

トレース画面



株式会社 東陽テクニカ

IB Trainer ×4エキササイザ

- **システムやデバイス・レベルの検証のために再現可能な任意の** InfiniBand 4x**パケットを生成**
- Invariant CRC / Variant CRC、**パケットヘッダ、パケットフレーミング、ランニングディスパリティ、**8B/10B**シンボルなどのエラーをトラフィックストリームにインジェクション可能**
- InfiniBand**コンポーネントの機能テストのため、リンク・トレーニング、アイドルデータの送信、スキップオーダードセット、リンクレベルのフロー制御が可能**

株式会社 東陽テクニカ

Infiniband SPEC1.1のサポート

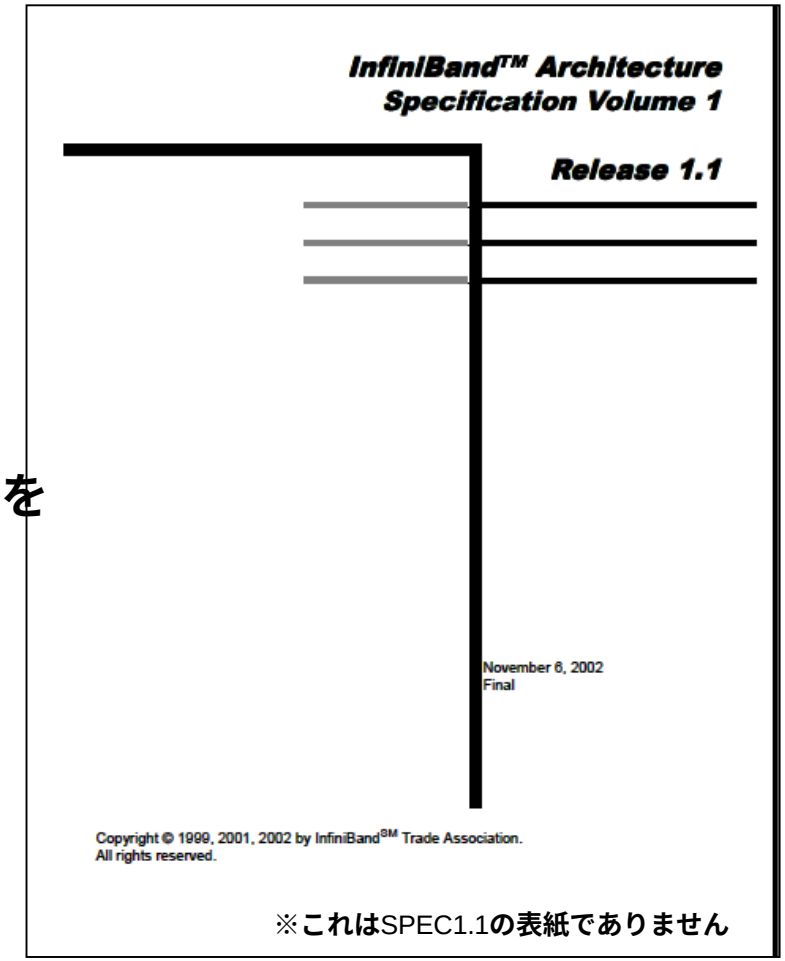
MAD (Management Datagram)のアーキテクチャ変更を含む InfiniBand SPEC1.1をサポート

以下の新しいMADパケットを デコード可能

- Subnet Management Class
- Subnet Administration Class
- Communication Management Class

以下の新しいトランザクションタイプを デコード可能

- Reliable Multi-Packet Transaction Protocol (RMPP) for SADM



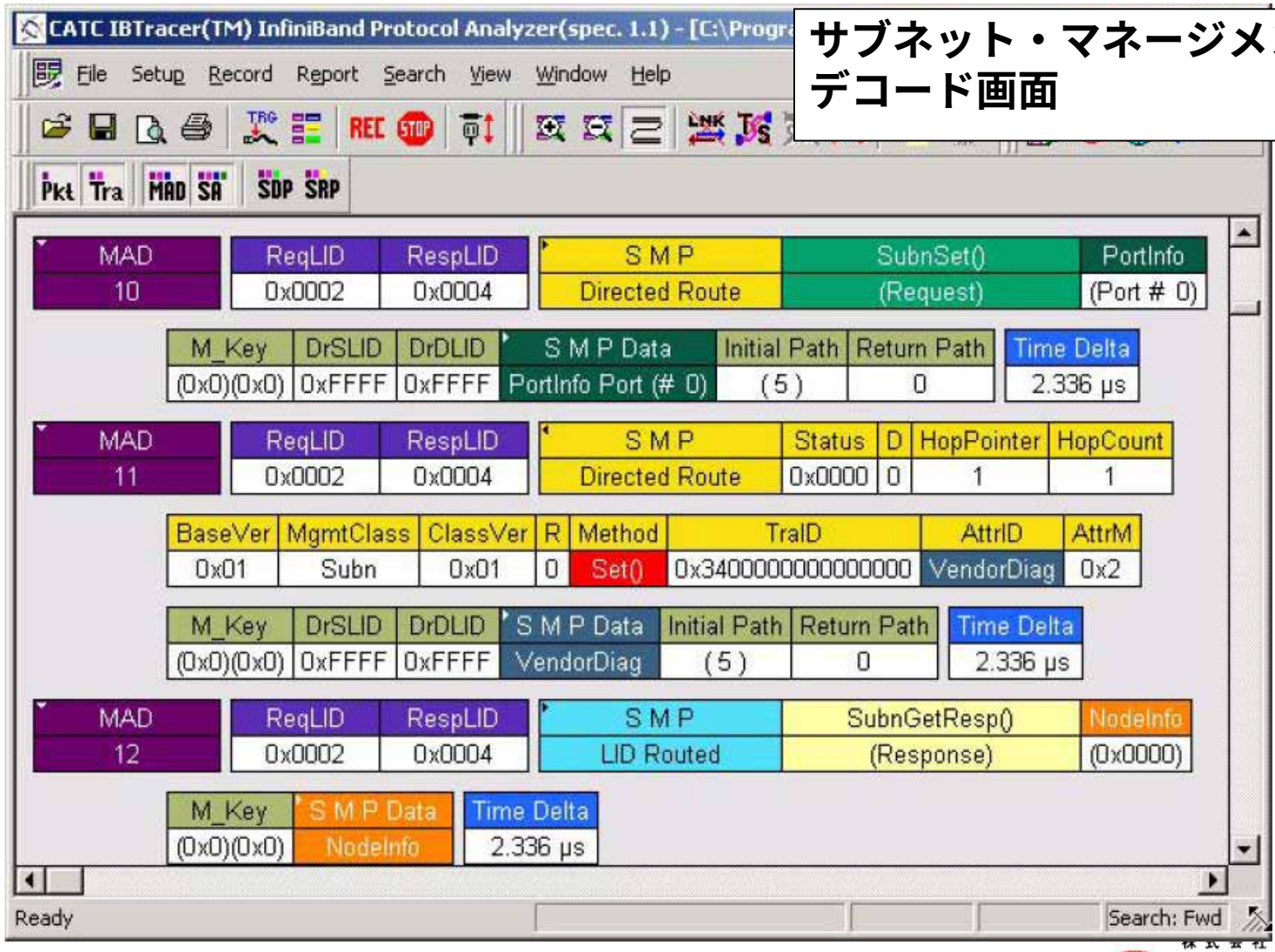
※これはSPEC1.1の表紙でありません

株式会社 東陽テクニカ

SPEC1.1：Subnet Management Class (1)

- サブネット・マネージメントクラスは、サブネット内の構成、モニタリング、ノードのクエリーを提供する。
- マスター・サブネット・マネジャーは、IBサブネットを初期化し構成する重要なエレメントであり、サブネットの初期化プロセスの一部として選択される。

SPEC1.1：Subnet Management Class (2)



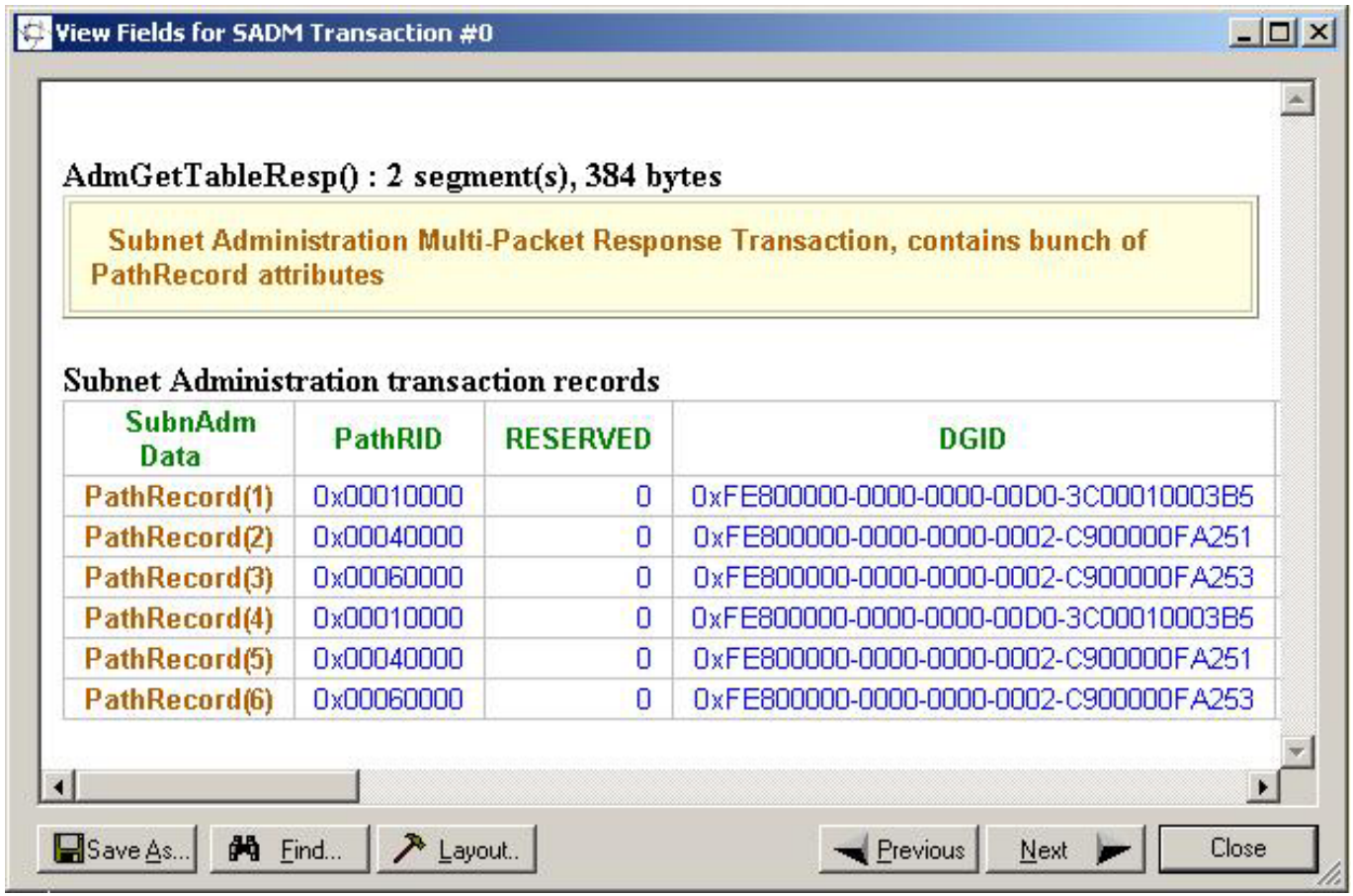
サブネット・マネージメントのデコード画面

東陽テクニカ

SPEC1.1：Subnet Administration Class (1)

- **サブネット・アドミニストレーション (SA) クラスは、サブネットの運営に必要な情報の格納場所を提供する。**
- **サブネット・トポロジとパーティション情報、データ・パス、イベント、サービスレベルの通知を含む。**

SPEC1.1：Subnet Administration Class (2)

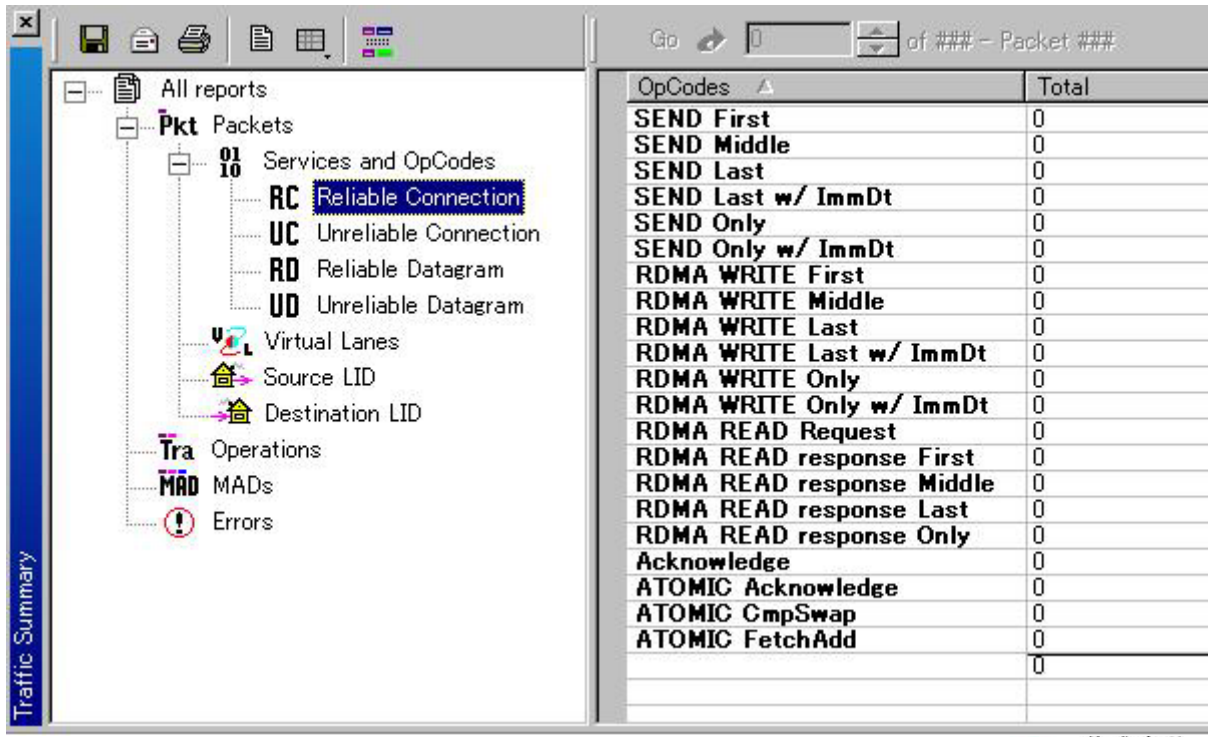


SubnAdm Data	PathRID	RESERVED	DGID
PathRecord(1)	0x00010000	0	0xFE800000-0000-0000-00D0-3C00010003B5
PathRecord(2)	0x00040000	0	0xFE800000-0000-0000-0002-C900000FA251
PathRecord(3)	0x00060000	0	0xFE800000-0000-0000-0002-C900000FA253
PathRecord(4)	0x00010000	0	0xFE800000-0000-0000-00D0-3C00010003B5
PathRecord(5)	0x00040000	0	0xFE800000-0000-0000-0002-C900000FA251
PathRecord(6)	0x00060000	0	0xFE800000-0000-0000-0002-C900000FA253

株式会社 東陽テクニカ

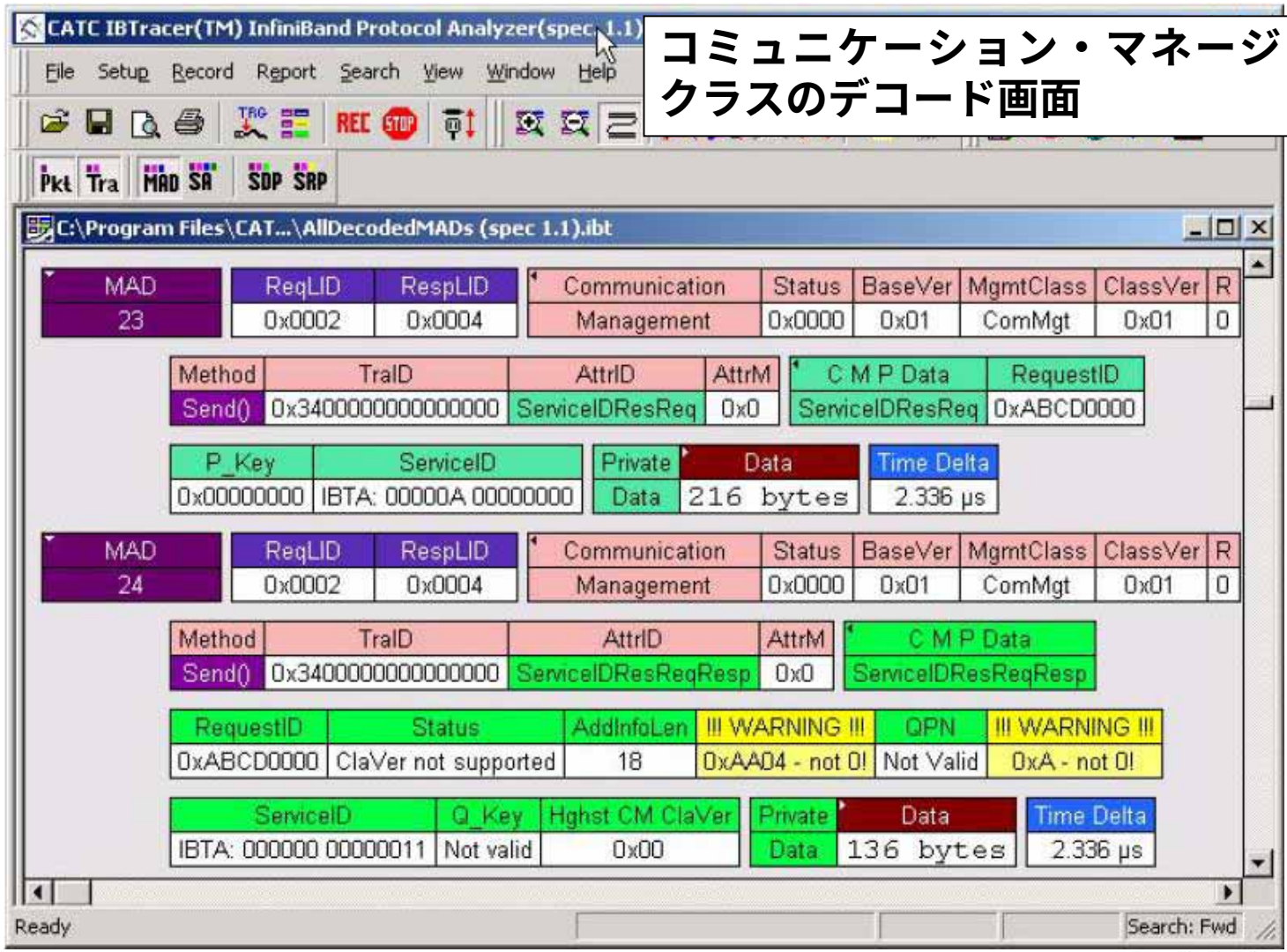
SPEC1.1：Communication Management Class (1)

- **コミュニケーション･マネージメント** (CM) **クラスは、** IB Reliable Connection、Unreliable Connection、Reliable Datagram **転送サービスのタイプのためにチャネルの確立、維持、開放に使用されるメカニズムを含む。**



株式会社 東陽テクニカ

SPEC1.1：Communication Management Class (2)



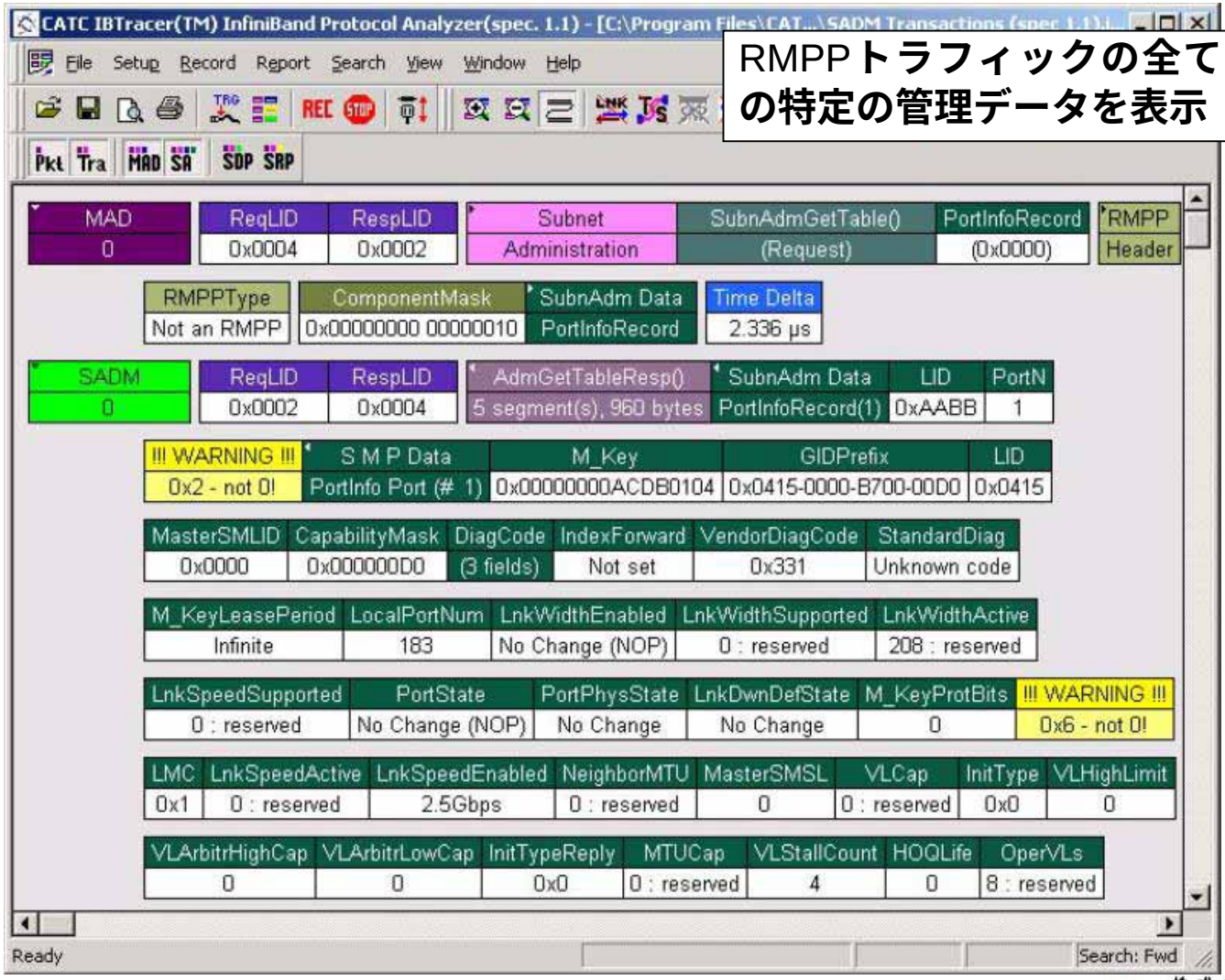
コミュニケーション・マネージメントクラスのデコード画面

株式会社 東陽テクニカ

SPEC1.1：Reliable Multi-Packet Transaction Protocol RMPP (1)

- InfiniBandアーキテクチャでは、マネージメントクラスがシングルのMADより大きい量のデータを確実に転送しなければならない状況がある。
- これを円滑に行うために、SPEC1.1ではシングルの論理トランザクションで大量のデータ（最高2^{32}パケット）を転送可能なRMPPを定めた。

SPEC1.1：Reliable Multi-Packet Transaction Protocol (RMPP) for SADM (2)



RMPPトラフィックの全てのポートの特定の管理データを表示

株式会社 東陽テクニカ

Lane Reversalについて

- The IBTracer 4X アナライザは、IBポートの Lane Reversalをサポート
- IBA 4Xリンクで接続するときに起こりうる、正しくないピンアサインを補正をオンチップ・ロジックで行うことを目的としたInfiniBandの特長である
- Lane Reversalを実装したIB 製品は、反転したレーン番号の順序を使用して4x linksの接続をするためPWB (printed wiring board) レイアウト を割り当てる
- IB Tracer 4X は、自動的に反転したレーンコンディションを検出するためにトレーニングシーケンス(TS1 or TS2)の間に送信されたレーン番号シンボルを使用する
- IB Tracer は、レーンの物理構成に関係なく論理的にIBトラフィックの表示、トリガリング、記録が可能

Script Verification Engineについて

- InfiniBand Trade Association でプラグテスト用にデザインされたコンプライアンス･テストスクリプトをサポート
- ユーザ独自のテストスクリプトの生成可能
- トレースデータからテストスクリプトの生成可能
- Lane-Reversalモードなどの拡張機能もサポート
- プラグテストにおいて
  - Infiniband TAの複雑なリンクレイヤテスト(c1v07-053)を自動化
    - リンクパケットが正しいオペコード を使用しているか (リンクの初期化の間、アーム、アクティブステート)
    - リンクパケット間の平均インターバル時間
    - パケットレングスフィールドの正確な計算

東陽テクニカ　ストレージアプリケーション用 プロトコル解析評価ツール製品ライン

Infiniband　アナライザ･ジェネレータ

Serial Attached SCSI　アナライザ･ジェネレータ

Serial ATA　アナライザ･ジェネレータ

ATA/ATAPI　アナライザ＆テスター･ジェネレータ

Ultra320SCSI　アナライザ＆テスタ

iSCSI / Fibre Channel　アナライザ＆テスタ

PCI Express　アナライザ＆ジェネレータ

I^2C　アナライザ＆ジェネレータ

株式会社 東陽テクニカ