



Agilent Technologies

IP-SAN Overview

2004/10/1

produced by:

**導体事業本部
前田 明德**

presented by:



**ネットマークス
小助川 通**

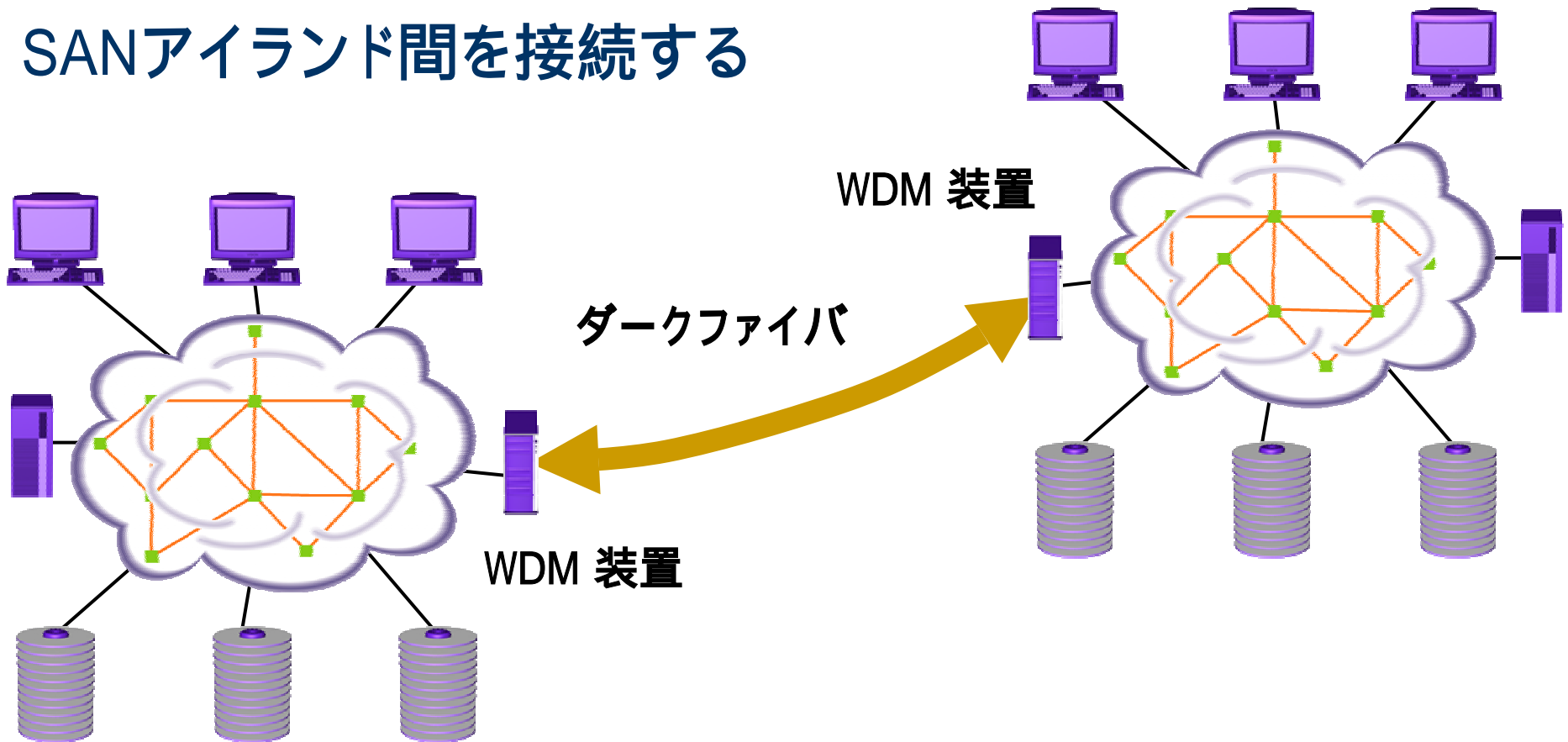
IP-SAN のアプリケーション

- **SAN 間の接続**
 - FCIP
- **IP ファブリックを用いた接続**
 - iFCP
- **ファイバチャネルの完全置き換え**
 - iSCSI



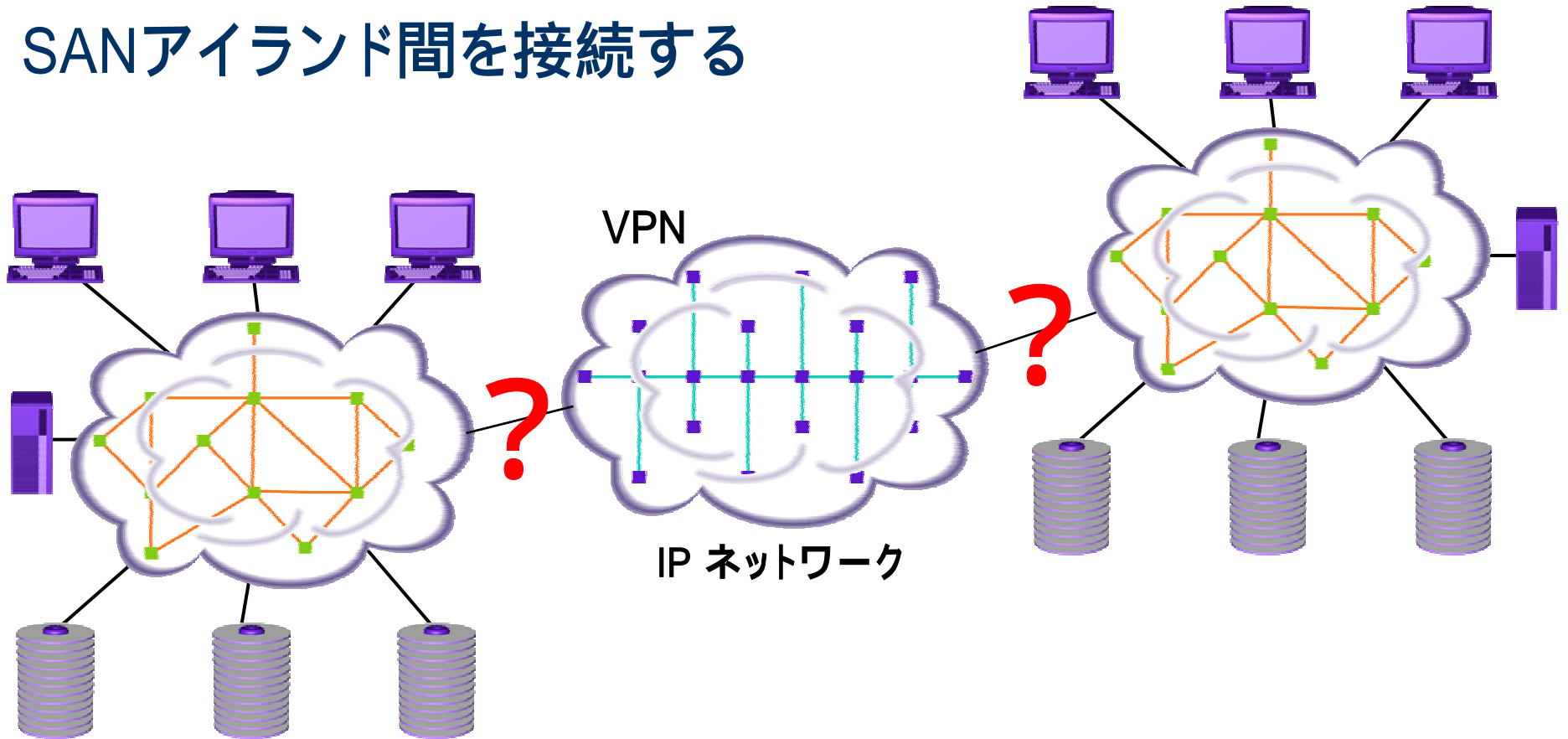
SAN 間の接続

SANアイランド間を接続する



SAN 間の接続 (IP ネットワーク)

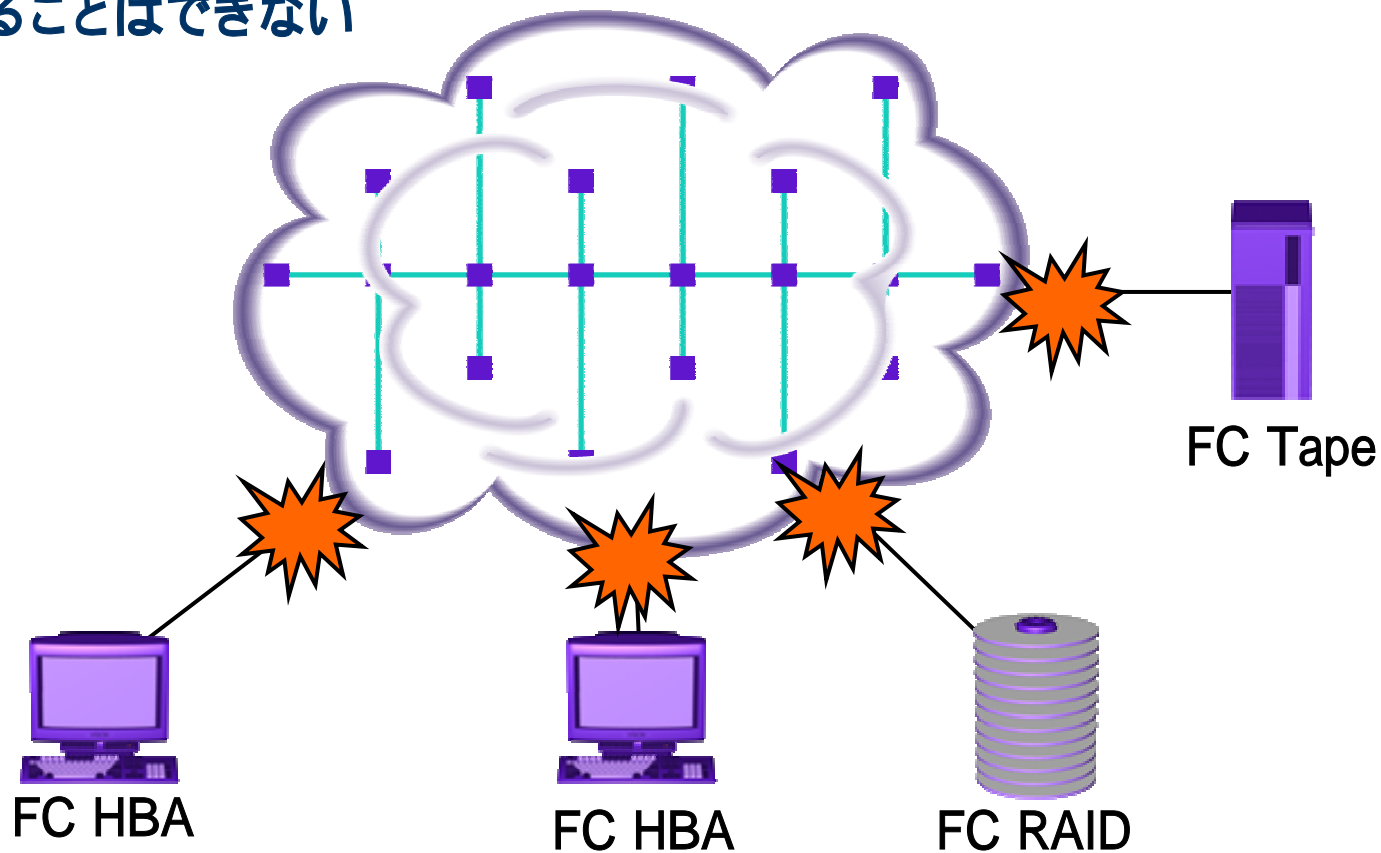
SANアイランド間を接続する



IP ファブリックを用いた接続

FC 装置を IP Network に
直接接続することはできない

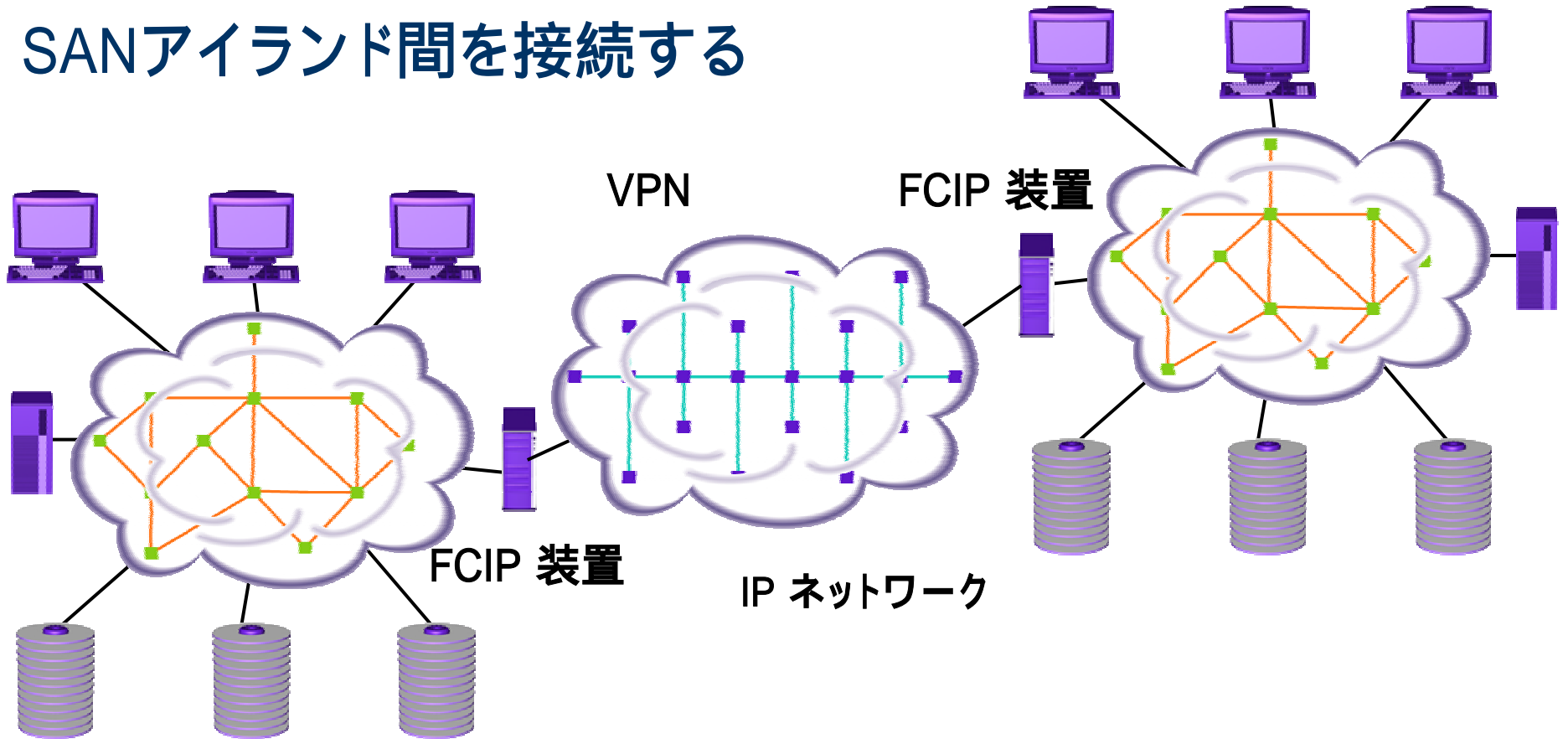
IP ネットワーク



Agilent Technologies

SAN 間の接続 (FCIP)

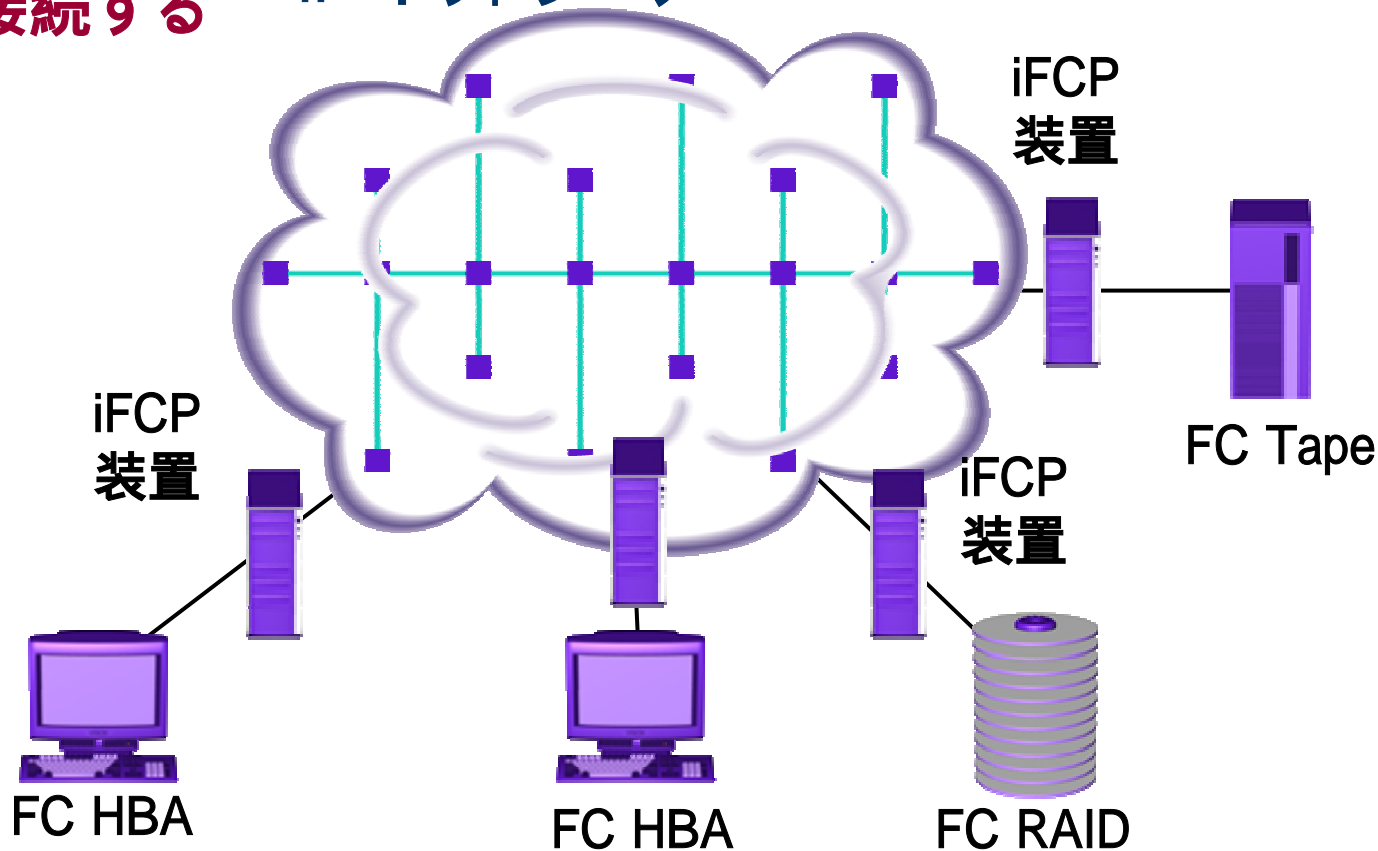
SANアイランド間を接続する



IP ファブリックを用いた接続

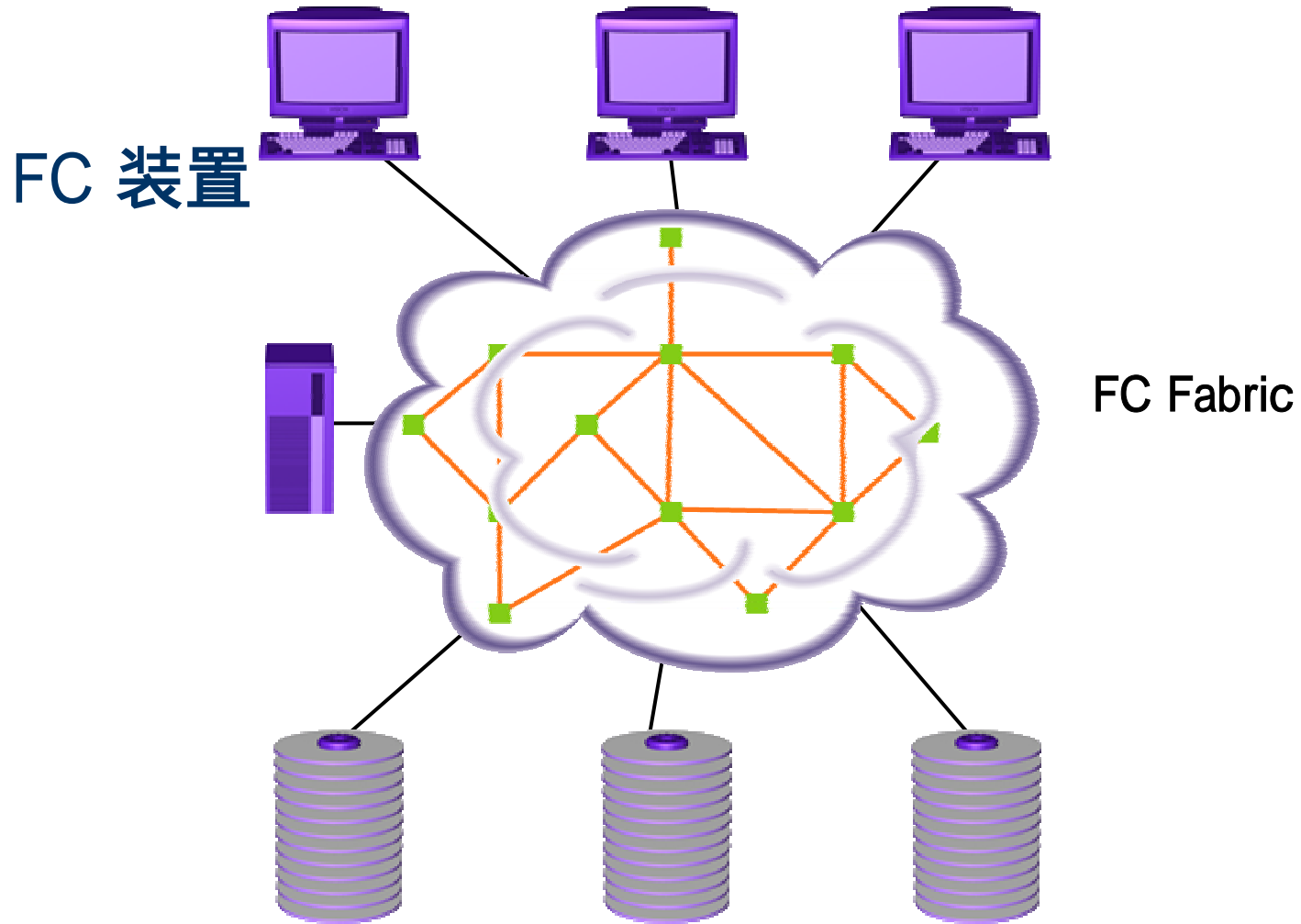
iFCP 装置を
使用して接続する

IP ネットワーク

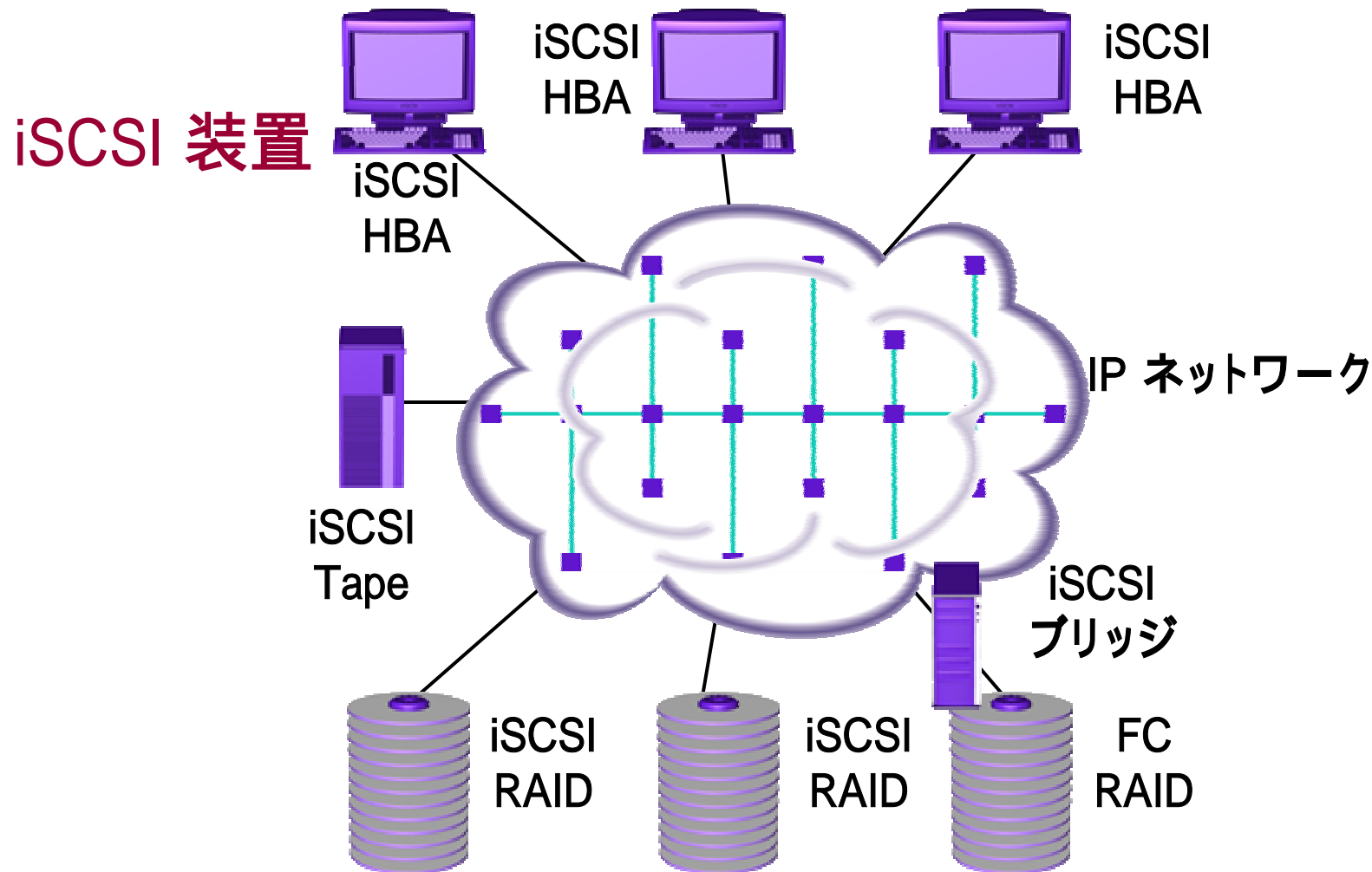


Agilent Technologies

ファイバチャネルの完全置き換え



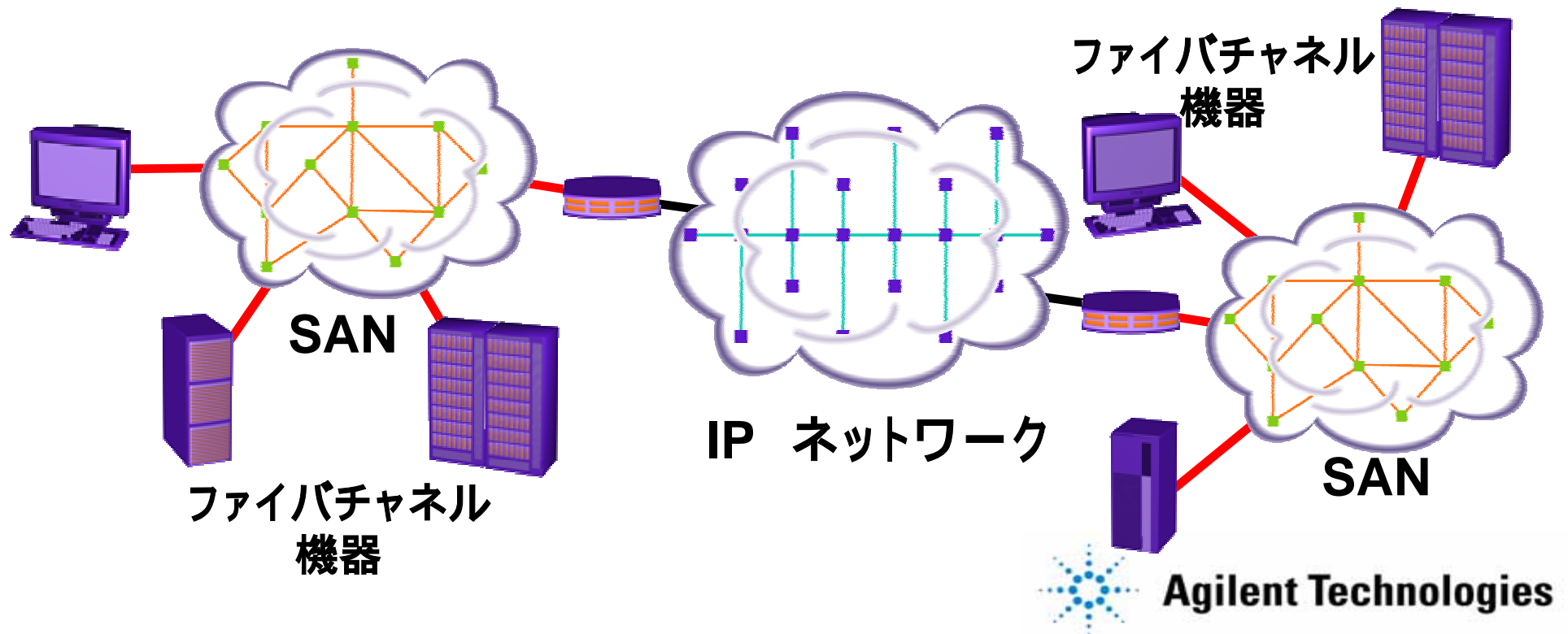
ファイバチャネルの完全置き換え



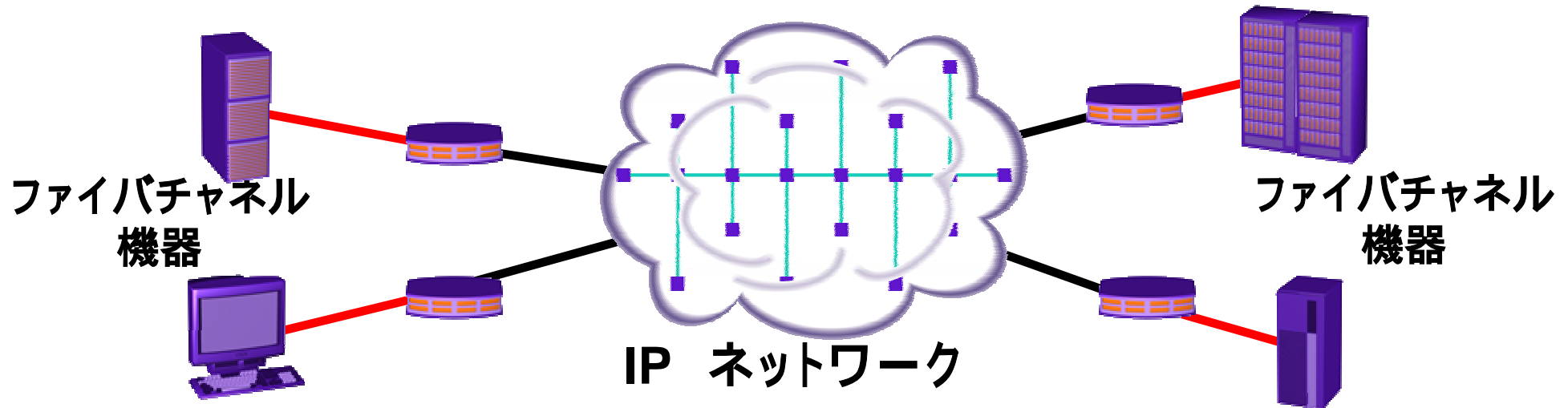
FCIP



- IETF ips-WG で標準化が完了 RFC3821
- FCフレームを IP ネットワークを通して送る
- IPネットワークを伝送路としたFCリンク(1対1)を延長
- 専用の機器で SAN と IP ネットワークをつなぐ



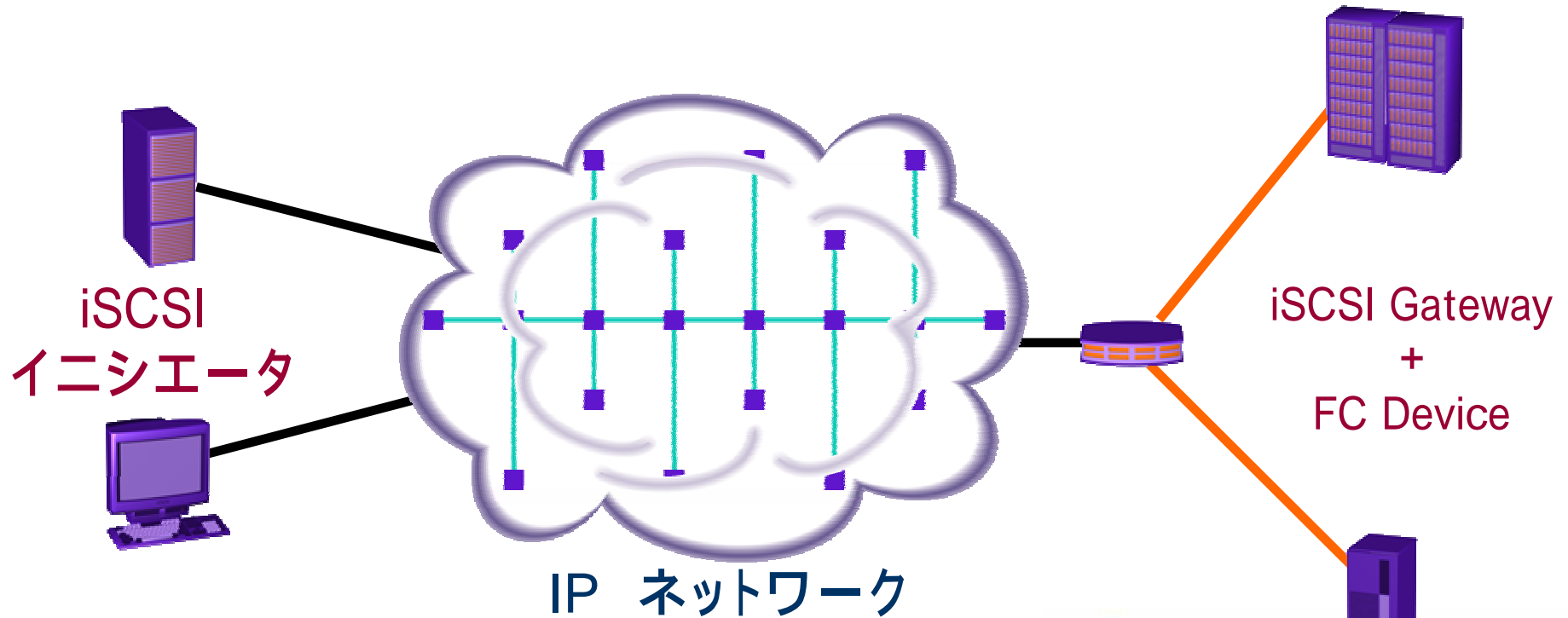
- IETF ips-WG で標準化 (draft-ietf-ips-ifcp-14)
- FCフレームを IP ネットワークを通して送る
- IPネットワークを用いたFCファブリック(1対n)の拡張
- 専用の機器でFC機器 と IP ネットワークをつなぐ



iSCSI



- IETF ips-WG で標準化が完了 RFC3720
- SCSI データを TCP/IP パケットで送る
- イニシエータ : iSCSI 対応の NIC/HBA が必要
- ディスク/テープ : iSCSI に対応することが必要





Agilent Technologies

iSCSI

2004/10/1

presented by:

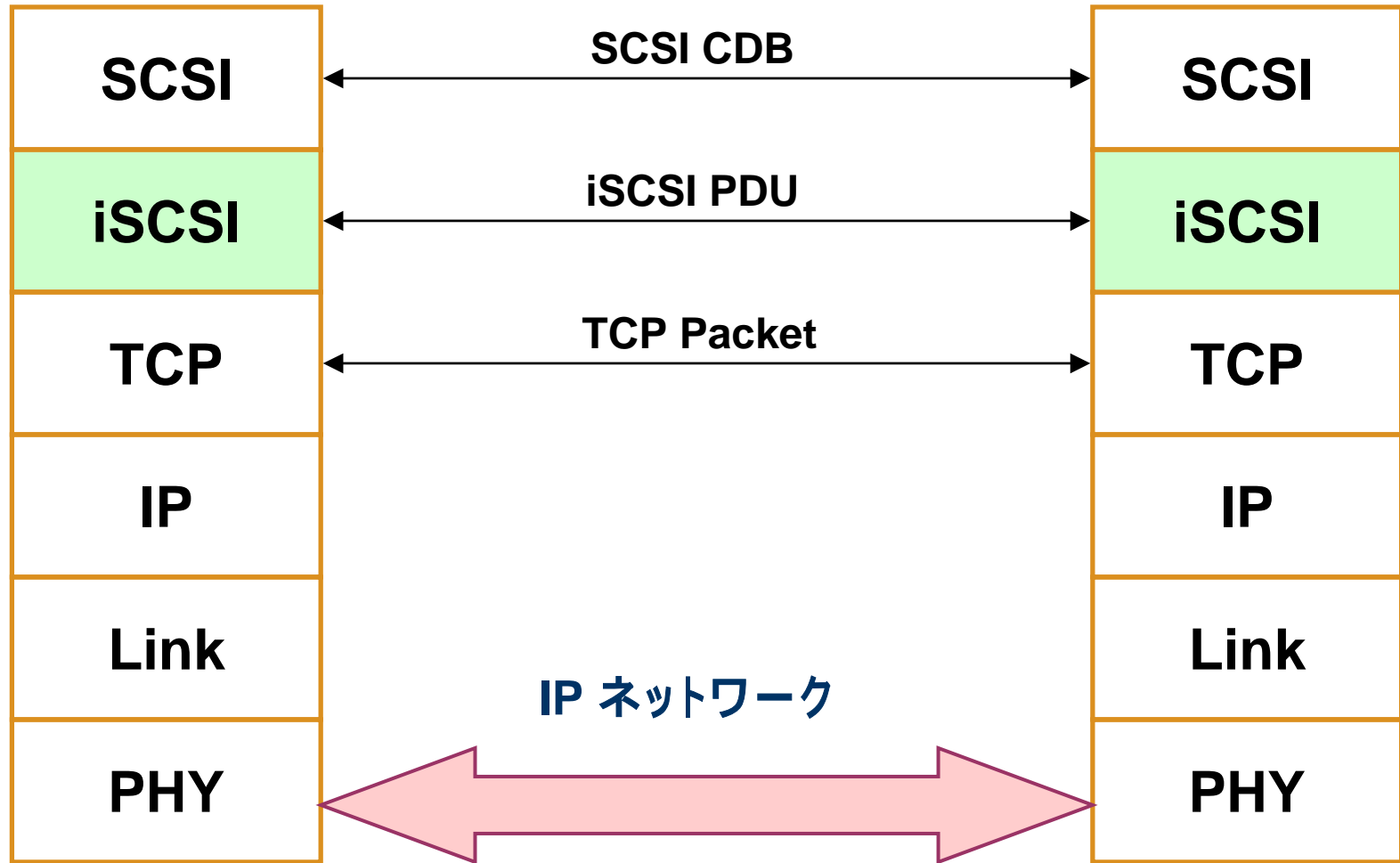
**半導体事業本部
前田明德**

iSCSI

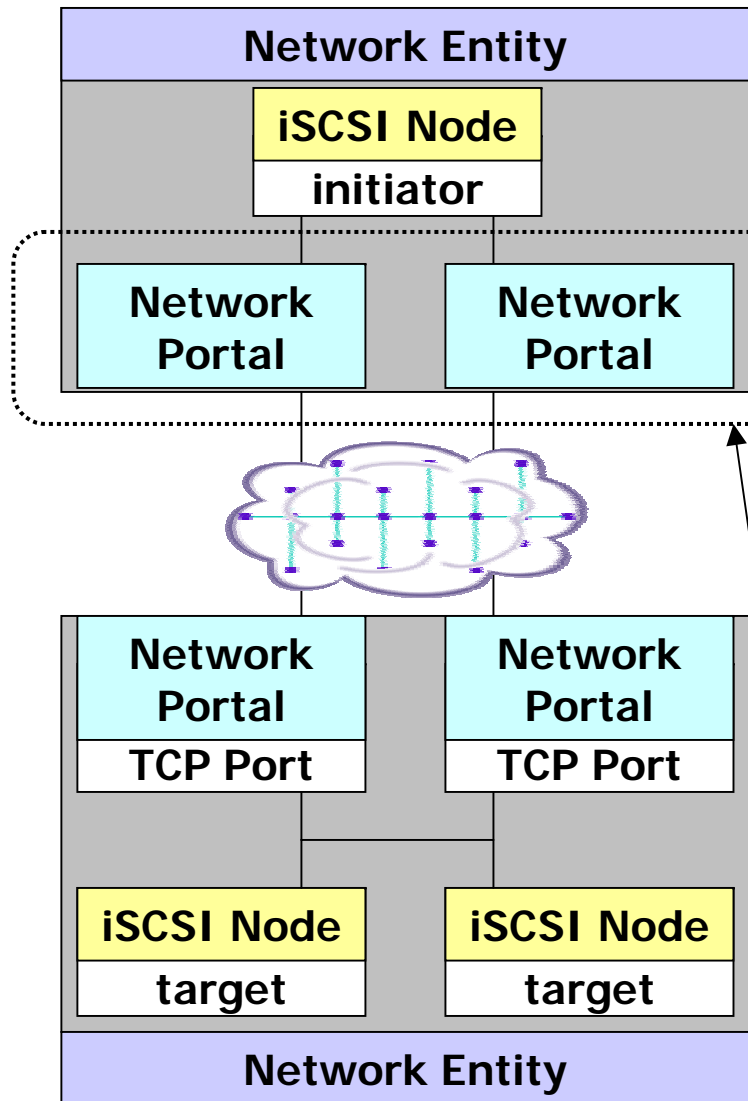
- internet Small Computer System Interface の略
- SCSI over TCP/IP
- SCSI データをカプセル化して TCP で送る
- Fibre Channel とはまったく独立の規格
- Fibre Channel の SCSI-FCP と同等
- SCSI Architecture Model - 2 (SAM-2)に対応



iSCSI プロトコル・モデル



iSCSI アーキテクチャー



- **Network Entity**

- IP でアクセス可能なデバイスまたはゲートウェイ。
- 1つ以上の iSCSI Node と Network Portal を持つ。

- **iSCSI Node**

- iSCSI initiator または target の実体。
- 各ノードは固有の iSCSI Name を持つ。

- **Network Portal**

- iSCSI Node と IP Network の玄関口。
- iSCSI Node と Network Portal は多対多。
- iSCSI initiator の場合 IP Address のみ。
- iSCSI target の場合 IP Address + TCP Port 番号を持つ。

- **Portal Group**

- マルチコネクションサポート。



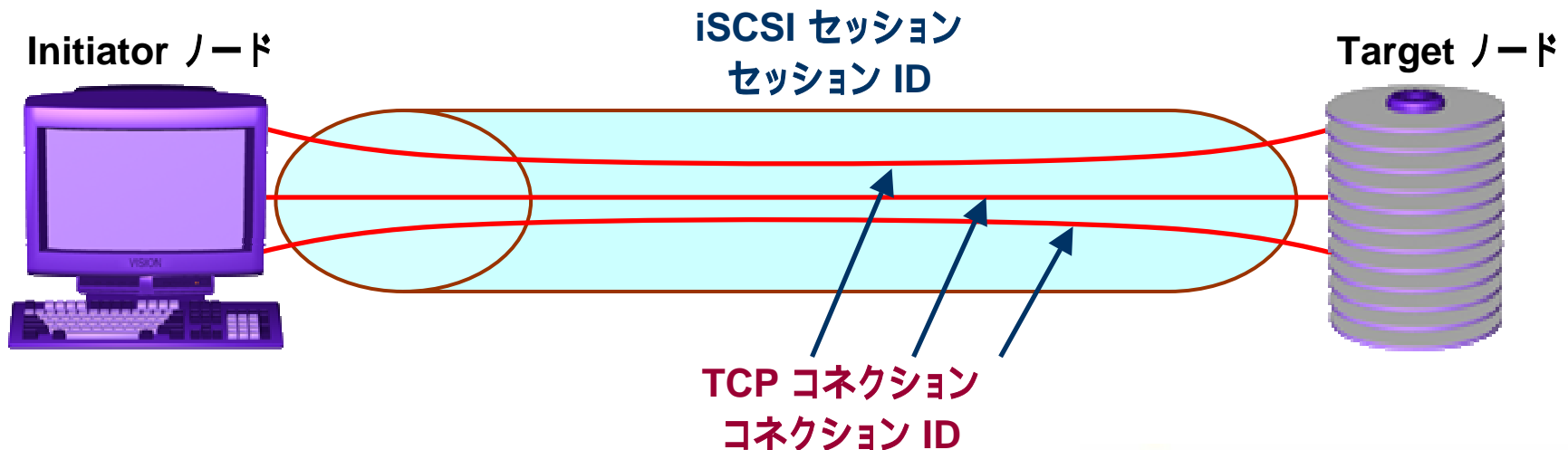
iSCSI ネーム

- iSCSI ノードは iSCSI ネームを持つ
 - “iqn” ネーム: iSCSI Qualified Name
 - iqn.<Date>.<Domain Name>[:Identifier]
 - “iqn.2001-04.com.agilent:ramdisk”
 - “eui” ネーム: IEEE EUI-64 Name
 - eui.<EUI-64 identifier>
 - “eui.0123456789ABCDEF”



iSCSI セッションと TCP コネクション

- iSCSI ノードは iSCSI セッションを通して他のノードと通信を行う
 - セッションにはセッション ID が割り当てられる
 - ISID: Initiator Session ID
 - TSIH: Target Session Identifier Handler
- iSCSI セッションは複数の TCP コネクションを持つことが可能
 - 各 TCP コネクションにはコネクション ID が割り当てられる



iSCSI PDU

Fibre Channel FCP

- FCP_CMND
- FCP_XFER_RDY

- FCP_DATA
- FCP_RSP

iSCSI PDU

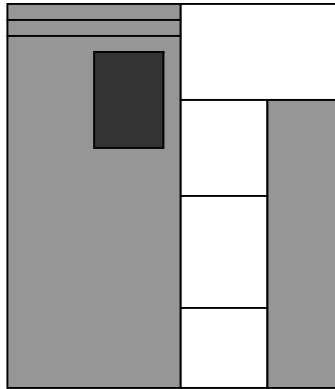
- SCSI Command
- Ready To Transfer (R2T)

- SCSI Data Out/In
- SCSI Response

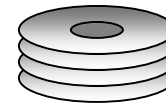


Read Operation

iSCSI Initiator



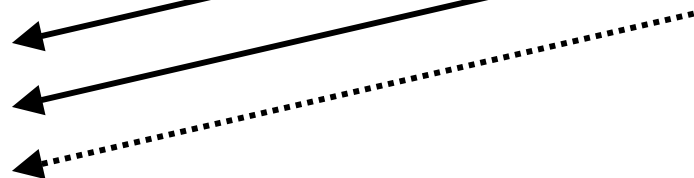
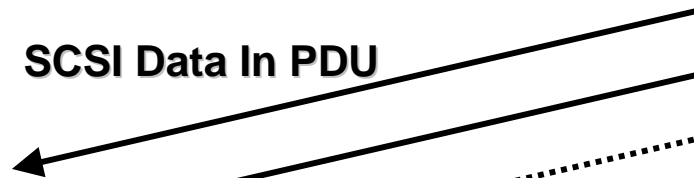
iSCSI Target



SCSI Command PDU



SCSI Data In PDU



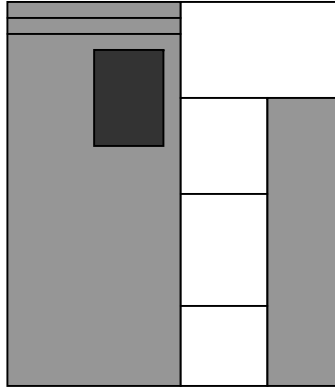
SCSI Response PDU



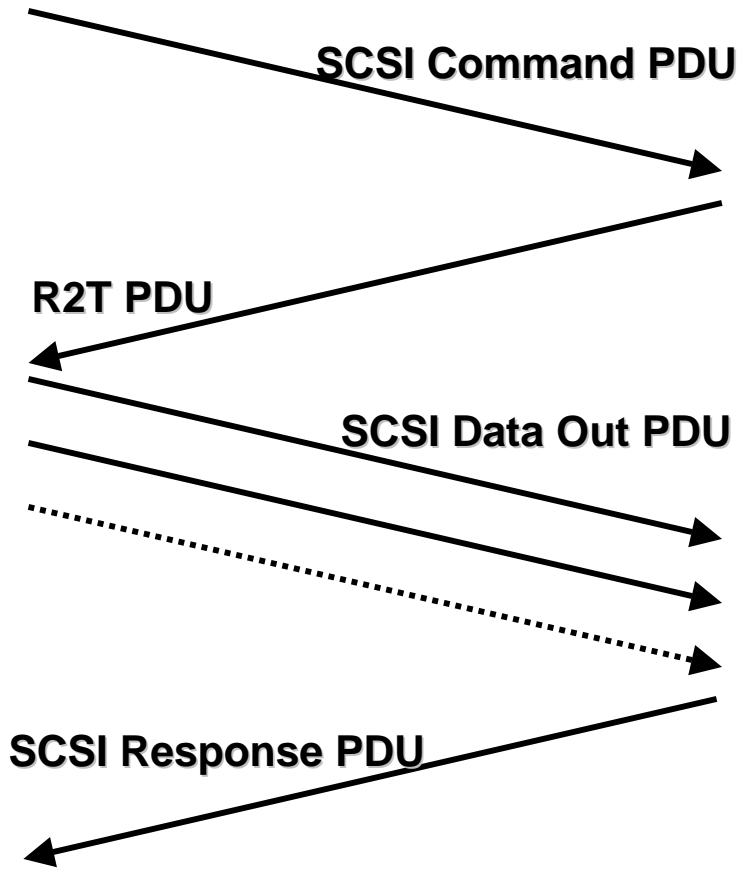
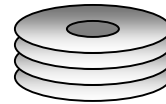
Agilent Technologies

Write Operation

iSCSI Initiator



iSCSI Target



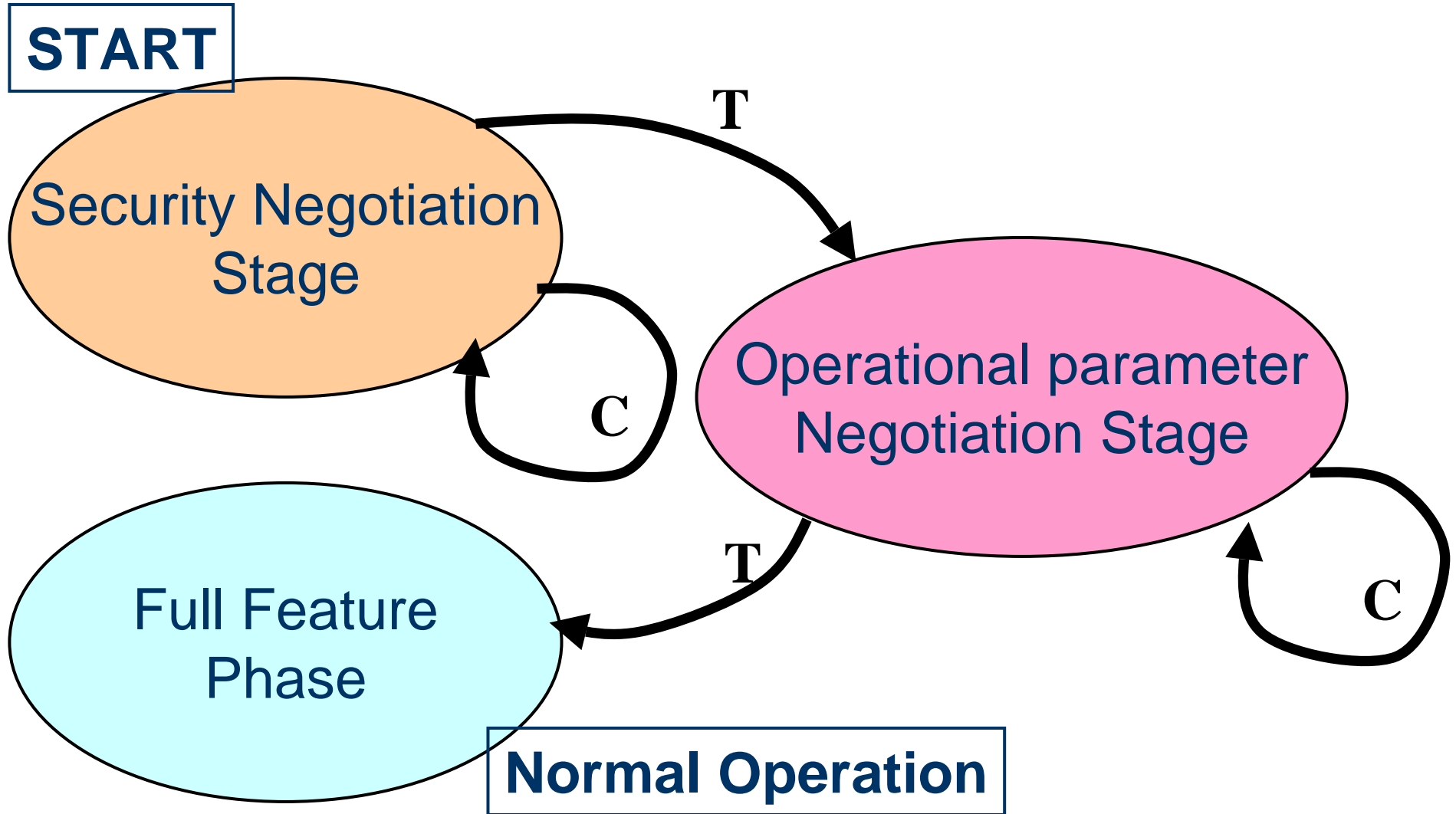
Agilent Technologies

iSCSI ログイン

- ログインにより、TCP コネクションを iSCSI で使用できるようにする
- ログインに先立ち IPsec による TCP コネクションをセキュア化する (オプション)
- Initiator は Login Request PDU を Target に送る
- Target は Login Response PDU を返信する
- ログインには二つのステージがある
 - Security Negotiation Stage
 - Operational Parameter Negotiation Stage
 - Login Request PDU と Login Response PDU の T(Transfer), C(Continue) ビットでステージを移動する



iSCSI ログイン ステージ遷移



ログイン・ネゴシエーション(1)

- ログインでは Text Operational Key を使用してノードの特徴を知らせると共に、通信パラメタをネゴシエーションする
 - “TargetName=iqn.2001.04.com.agilent:ramdisk”
 - “MaxConnections=4”
 - “InitialR2T=Yes”
 - など



ログイン・ネゴシエーション(2)

- 一方が Text Operational Key を送り、もう一方が応答する
 - リストネゴシエーション
 - Request: “AuthMethod=KRB5,SRP,None”
 - Response: “AuthMethod=SRP”
 - シンプル・バリュー・ネゴシエーション
 - Request: “InitialR2T=Yes”
 - Response: “InitialR2T=No”
- 選択できない時などには以下のように返すことがある
 - “None”, “Reject”, “Irrelevant”, “NotUnderstood”
- 1つのText Operational Key は2回ネゴシエーションできない



ログインKEY



ステージ	KEY の名前	送信者	ステージ	KEY の名前	送信者
オペレーショナル パラメタ ネゴシエーション ステージ	HeaderDigest	イニシエータ ターゲット	セキュリティ ネゴシエーション ステージ	SessionType	イニシエータ
	DataDigest	イニシエータ ターゲット		InitiatorNmae	イニシエータ
	MaxConnections	イニシエータ ターゲット		TargetName	イニシエータ
	TargetName	イニシエータ		TargetAddress	ターゲット
	InitiatorName	イニシエータ		InitiatorAlias	イニシエータ
	TargetAlias	ターゲット		TargetAlias	ターゲット
	InitiatorAlias	イニシエータ		TargetPortalGroypTag	ターゲット
	TargetAddress	ターゲット		AuthMethod	イニシエータ ターゲット
	TargetPortalGroypTag	ターゲット			
	InitialR2T	イニシエータ ターゲット	フル フィーチャ フェーズ	SendTargets	イニシエータ
	ImmediateData	イニシエータ ターゲット		TargetName	ターゲット
	MaxRecvDataSegment Length	イニシエータ ターゲット		TargetAlias	ターゲット
	MaxBurstLength	イニシエータ ターゲット		InitiatorAlias	イニシエータ
	FirstBurstLebgh	イニシエータ ターゲット		TargetAddress	ターゲット
	DefaultTime2wait	イニシエータ ターゲット		MaxRecvDataSegment Length	イニシエータ ターゲット
	DefaultTime2Retain	イニシエータ ターゲット			
	MaxOutstandingR2T	イニシエータ ターゲット			
	DataPDUInOrder	イニシエータ ターゲット			
	DataSequenceInOrder	イニシエータ ターゲット			
	ErrorRecoveryLevel	イニシエータ ターゲット			
SessionType	イニシエータ				



ログインの終了

- エラーがあったら、そこでリジェクトする
 - Reject PDU が送られる
- ログインが成功したら、Full Feature フェーズに移行してログインを終了する
 - Full Feature フェーズでは、通常の iSCSI の通信を行うことができる



ログアウト

- 通信が終了したら Logout Request PDU を送り、ログアウトを要求する
 - Reason Code
 - 0: Close the session
 - 1: Close the connection
 - 2: Remove the connection for recovery
- Logout Response PDU でログアウトのステータスを返信する
 - Logout response
 - 0: Closed successfully
 - 1: Connection ID not found
 - 2: Fail to close



SCSI Command フレーム

	Byte 0								Byte 1								Byte 2								Byte 3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0		I	0x01						F	R	W	0	0	ATTR				Reservrd														
4	Total AHS Length								Data Segment Length																							
8	Logical Unit Number (LUN)																															
12																																
16	Initiator Task Tag																															
20	Expected Data Transfer Length																															
24	Command Sequence Number																															
28	Expected Status Sequence Number																															
32	Command Data Block (CDB)																															
36																																
40																																
44																																
48	Header Digest																															
52	Data Area																															
...																																



SCSI DATA-out フレーム

	Byte 0								Byte 1								Byte 2								Byte 3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0			0x05						F	Reserved																						
4	Total AHS Length								Data Segment Length																							
8	Logical Unit Number (LUN) or Reserved																															
12																																
16																																
16	Initiator Task Tag																															
20	Target Transfer tag																															
24	Reserved																															
28	Expected Status Sequence Number																															
32	Reserved																															
36	Data Sequence Number																															
40	Buffer Offset																															
44	Reserved																															
48	Header Digest																															
52	Data Area																															
...																																



SCSI DATA-in フレーム

	Byte 0								Byte 1								Byte 2								Byte 3								
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	
0			0x25						F	A	0	0	0	0	O	U	S	Reserved								Status or Reserved							
4	Total AHS Length								Data Segment Length																								
8	Logical Unit Number (LUN) or Reserved																																
12																																	
16	Initiator Task Tag																																
20	Target Transfer tag																																
24	Status Sequence Number or Reserved																																
28	Expected Command Sequence Number																																
32	Maximum Command Sequence Number																																
36	Data Sequence Number																																
40	Buffer Offset																																
44	Residual Count																																
48	Header Digest																																
52	Data Area																																
...																																	



SCSI Response フレーム

	Byte 0								Byte 1								Byte 2								Byte 3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0			0x21						1			0	u	O	U		Response								Status							
4	Total AHS Length								Data Segment Length																							
8	Reserved																															
12	Reserved																															
16	Initiator Task Tag																															
20	Reserved																															
24	Status Sequence Number																															
28	Expected Command Sequence Number																															
32	Maximum Command Sequence Number																															
36	Expected Data Sequence Number or Reserved																															
40	Bidirectional Read Residual Count																															
44	Residual Count																															
48	Header Digest																															
52	Data Area																															
...																																





- HBA
 - Gbit Ethernet NIC + iSCSI ドライバ
 - TCP Offload Engine + iSCSI ドライバ
 - iSCSI HBA
- iSCSI ブリッジ
 - FC – iSCSI ブリッジ
- iSCSI ストレージ
 - 各社 iSCSI 対応のオプションは用意



- ありがとうございます