(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6135403号
(P6135403)

(45)発行日　平成29年5月31日(2017.5.31)　(24)登録日　平成29年5月12日(2017.5.12)

(51)Int.Cl.　　F I
G06F 11/07 (2006.01)　　G06F 11/07 172
G06F 11/34 (2006.01)　　G06F 11/07 140A
　　G06F 11/34 176

請求項の数 5　(全 27 頁)

(21)出願番号　特願2013-175250 (P2013-175250)
(22)出願日　平成25年8月27日 (2013.8.27)
(65)公開番号　特開2015-45905 (P2015-45905A)
(43)公開日　平成27年3月12日 (2015.3.12)
審査請求日　平成28年5月10日 (2016.5.10)

(73)特許権者　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100094525
弁理士　土井　健二
(74)代理人　100094514
弁理士　林　恒徳
(72)発明者　結城　和博
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

審査官　石川　亮

最終頁に続く

(54)【発明の名称】情報処理システム、情報処理システムの障害処理方法

(57)【特許請求の範囲】
【請求項１】
複数のノード間でメモリを共有する情報処理システムにおいて、
前記ノードの各々は、
複数の機能回路と前記機能回路を制御する制御装置と、
前記複数の機能回路から発生する割り込み要因を格納するレジスタとを有し、
前記複数のノードのうちの１のノードの前記制御装置は、
他の前記ノードの割り込み要因の発生に応じて前記レジスタの前記割り込み要因を受信し、前記割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、前記障害ノードの前記メモリへのアクセスを抑止後、前記他のノードから受信したログ情報に基づいて前記障害ノードの切り離し制御を行う情報処理システム。
【請求項２】
請求項１において、
前記１のノードは、網結合装置を備え、
前記他のノードは、データ処理を実行し、前記網結合装置を介して前記メモリにアクセスする処理装置を備える情報処理システム。
【請求項３】
請求項１または２において、
前記１のノードの前記制御装置は、前記障害として検出すべき割り込み要因の波及元と

なる割り込み要因が発生しているか否かを判定し、発生していない場合に、前記割り込み要因に対応するノードを前記障害ノードとして特定し、発生している場合に、前記波及元となる割り込み要因に対応するノードを前記障害ノードとして特定する情報処理システム。

【請求項４】

請求項３において、

前記１のノードは、

前記割り込み要因と、前記割り込み要因の波及元となる割り込み要因との対応関係を有する定義テーブルを有し、

前記１のノードの前記制御装置は、前記定義テーブルに基づいて、前記障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定する情報処理システム。

【請求項５】

複数のノード間でメモリを共有する情報処理システムの障害処理方法において、

前記ノードの各々は、

複数の機能回路と前記機能回路を制御する制御装置と、

前記複数の機能回路から発生する割り込み要因を格納するレジスタとを有し、

前記複数のノードのうちの１のノードの前記制御装置は、

他の前記ノードの割り込み要因の発生に応じて前記レジスタの前記割り込み要因を受信し、前記割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、前記障害ノードの前記メモリへのアクセスを抑止後、前記他のノードから受信したログ情報に基づいて前記障害ノードの切り離し制御を行う情報処理システムの障害処理方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、情報処理システム、情報処理システムの障害処理方法に関する。

【背景技術】

【０００２】

複数のノードを有する情報処理システムは、例えば、ビルディングブロック（BB：Building Block）構造を有する。例えば、複数のノードでメモリを共有する情報処理システムは、ノード間で、クロスバーを介してメモリを共有する。情報処理システムで動作するアプリケーションは、共有されたメモリを使用することによって、システムの処理性能の向上を図る。一方、それぞれのノードで動作するＯＳ（Operation system、以下、ＯＳと称する）やハイパーバイザ（hypervisor）は、各ノードのローカルメモリ上で動作する。ＯＳやハイパーバイザがローカルメモリ上で動作することにより、各ノードの独立性が高まり、システムの可用性が向上する。

【０００３】

このような情報処理システムにおいて、一部のノードのハードウェアに障害が発生した場合、障害が発生した障害ノードを検出すると共に、障害ノードをシステムから切り離した状態で、運用を再開することが求められる。ハードウェアの障害の検出は、例えば、特許文献１に記載される。

【先行技術文献】

【特許文献】

【０００４】

【特許文献１】特開２０１１-２４８６５３号公報

【発明の概要】

【発明が解決しようとする課題】

【０００５】

情報処理システムは、障害ノードの特定やシステムからの切り離しの要否を、障害の事

象を有するログ情報に基づいて、順次、解析する。したがって、情報処理システムのノード数や、障害の事象の種類の増加に伴って、障害ノードの特定やシステムからの切り離しの要否に係る解析時間も増加する。また、ログ情報のデータ量が膨大であることにより、ログ情報の収集にも時間を要する。

【０００６】

１つの側面は、本発明は、障害発生時に障害ノードによる他ノードの影響を早急に低減する情報処理システム、情報処理システムの障害処理方法を提供することを目的とする。

【課題を解決するための手段】

【０００７】

第１の側面は、複数のノード間でメモリを共有する情報処理システムにおいて、前記ノードの各々は、複数の機能回路と前記機能回路を制御する制御装置と、前記複数の機能回路から発生する割り込み要因を格納するレジスタとを有し、前記複数のノードのうちの１のノードの前記制御装置は、他の前記ノードの割り込み要因の発生に応じて前記レジスタの前記割り込み要因を受信し、前記割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、前記障害ノードの前記メモリへのアクセスを抑止後、前記他のノードから受信したログ情報に基づいて前記障害ノードの切り離し制御を行う。

【発明の効果】

【０００８】

第１の側面によれば、情報処理システムは、割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、障害ノードのメモリへのアクセスを抑止することで、障害発生時に障害ノードによる他ノードの影響を早急に低減する。

【図面の簡単な説明】

【０００９】

【図１】本実施の形態例における情報処理システムの概要を説明する図である。

【図２】図１の情報処理システムの構成の一例を示す図である。

【図３】図２に示したシステムボードの構成の一例を説明する図である。

【図４】図３のレジスタを説明する図である。

【図５】図１～図３で述べた、本実施の形態例における情報処理システムの一部のノードにおいて障害が発生した場合の処理の流れを説明する図である。

【図６】図５において述べた、マスターノードのシステム制御装置におけるログ情報の解析処理の概要を説明する図である。

【図７】図６のログ情報の解析、及び、ＦＮＬ解析処理に要する時間を例示する図である。

【図８】本実施の形態における情報処理システムの各ノードのソフトウェアモジュール図である。

【図９】図８において説明したＦＮＬ（Fail Node List）の一例を示す図である。

【図１０】割り込み要因が発生しＦＮＬが更新される間のマスターノードのシステム制御装置、及び、スレーブノードのシステム制御装置における、処理の流れを時系列に説明する図である。

【図１１】本実施の形態例におけるＦＮＬ解析部の処理、及び、ＦＮＬ更新部の処理を説明するフローチャート図である。

【図１２】波及先の割り込み要因の抑止処理を説明するフローチャート図である。

【図１３】ＦＮＤＢ（Fail Node DB）の一例を示す図である。

【図１４】アクション番号（ａｃｔ）を有する定義テーブルの具体例を示す図である。

【図１５】具体例におけるメモリのアクセスの抑止範囲を説明する図である。

【発明を実施するための形態】

【００１０】

以下、図面にしたがって本発明の実施の形態を説明する。ただし、本発明の技術的範囲

はこれらの実施の形態に限定されず、特許請求の範囲に記載された事項とその均等物まで及ぶものである。

【００１１】

［情報処理システムの概要］

図１は、本実施の形態例における情報処理システム１の概要を説明する図である。図１に示す情報処理システム１は、ＨＰＣ（High Performance Computing）モデル等の計算機システムである。このようなシステムは、ビルディングブロック（BB：Building Block）構造によって構成される。各ビルディングブロック１０ａ～１０ｅは、図１に示すシステムボード１Ａ～１Ｅを収容し、ラックに抜き差し可能である。また、図１の情報処理システム１は、複数のシステムボード１Ａ～１Ｅと、網結合装置（以下、クロスバスイッチと称する）２を備えるシステムボードとを有する。各システムボード１Ａ～１Ｅは、クロスバスイッチ２を介して、相互に接続する。なお、図１には、５つのシステムボード１Ａ～１Ｅが示されるが、情報処理システム１は、例えば、１６台のシステムボードを有する。

【００１２】

また、システムボード１Ａは、複数のＣＰＵ（Central Processing Unit）１２ａとメモリ３、１１ａと、Ｉ／Ｏ（Input Output）装置１３ａとを有する。また、メモリ３、１１ａの一部の領域は、情報処理システム１が有する全てのＣＰＵが共用する共有メモリ３として使用され、他の領域は、ＣＰＵ１２ａがカーネルデータ等を格納するローカル領域１１ａとして使用される。他のシステムボード１Ｂ～１Ｅも、システムボード１Ａと同様の構成を有する。以下、各システムボードをノードと称する。

【００１３】

また、ノード１Ａのファームウェア層１４ａでは、例えば、ハイパーバイザ（hypervisor）と呼ばれる制御ソフトウェアが動作する。ハイパーバイザは、ノード１Ａのリソースを論理的に分割して、１つまたは複数の論理パーティションＤａ、Ｄｂを生成する。複数の論理パーティションＤａ、Ｄｂが生成されることにより、１つのノード上で複数のＯＳ（Operation system、以下、ＯＳと称する）が動作可能になる。なお、図１の例において、各論理パーティションＤａ、Ｄｂ上で動作するＯＳ（例えば、Ｓｏｌａｒｉｓ（登録商標））は、異なる種類のＯＳであってもよい。

【００１４】

また、各論理パーティションＤａ～Ｄｈ上で動作するアプリケーションｐａ～ｐｈは、例えば、共有メモリ３を使用する。即ち、本実施の形態例では、各ノードが共有メモリ３の一部を有し、各ノードが他ノードの共有メモリ３を利用する分散型共有メモリを構成する。そして、アプリケーションｐａ～ｐｈは、共有メモリ３に記憶された共有の情報に基づいて、所定の処理を行う。また、ハイパーバイザやＯＳは、各々のローカルメモリ１１ａ～１１ｅ上で動作することにより独立性が高まり、システムの可用性が向上する。

【００１５】

分散型共有メモリ３を有する情報処理システム１において、例えば、ノード１ＡのＣＰＵ１２ａが、アプリケーションｐａの実行にあたり、共有メモリ３上のノード１Ａとは別のノード（例えば、ノード１Ｂ）の共有メモリ３の領域にアクセスする場合、ＣＰＵ１２ａは、クロスバスイッチ２を介して、ノード１Ｂの共有メモリ３の領域にアクセスのリクエストを送信する。また、ＣＰＵ１２ａが、自ノード１Ａの共有メモリ３の領域にアクセスする場合、直接接続を介して、メモリアクセスのリクエストを送信する。

【００１６】

［情報処理システムの構成］

図２は、図１の情報処理システム１の構成の一例を示す図である。図２において、図１で示したものと同一のものは、同一の記号で示す。図２に示すように、情報処理システム１は、例えば、処理装置としての１６台のシステムボード（ＳＢ：System Board）１Ａ～１Ｐと、４台のクロスバスイッチボックス２ＡＢ～２ＤＢとを有する。クロスバスイッチボックス２ＡＢ～２ＤＢがそれぞれ有するクロスバスイッチ２Ａ～２Ｄは、図１に示すクロスバスイッチ２に対応する。本実施の形態も、クロスバスイッチボックス２ＡＢ～２Ｄ

Ｂは、ビルディングブロック構造である。

【００１７】

図２の例において、クロスバスイッチボックス２ＡＢは、クロスバスイッチ２Ａと、システム制御装置（ＳＶＰ：SerVice Processor）Ｖ１とを有する。クロスバスイッチボックス２ＡＢのシステム制御装置Ｖ１は、クロスバスイッチ２Ａの状態監視、状態設定、及び、起動、停止制御等を行う。また、クロスバスイッチ２Ａは、スイッチ２ａとポートａｖ、ａｗ～ｄｖ、ｄｗ、ｑｖ、ｑｗ、ｒｖ、ｒｗ、ｓｖ、ｓｗとスイッチ２ａとを有する。スイッチ２ａは、通信経路を切り替える。他のクロスバスイッチボックス２ＢＢ～２ＤＢの構成も同様である。

【００１８】

また、図２の例において、それぞれのシステムボード１Ａは、２つのクロスバスイッチ２Ａとの接続用ポートａｘ、ａｙを有する。また、クロスバスイッチ２Ａも、各システムボード１Ａとの２つの接続用ポートａｖ、ａｗを有する。即ち、各システムボード１Ａは、２つの回線ｎ１、ｎ２によって、対応するクロスバスイッチ２Ａに接続する。このように、図２に示すクロスバスイッチ２Ａ～２Ｄは、接続対称との間に二重の回線を有する対称型のクロスバスイッチである。二重の回線を有するため、クロスバスイッチ２Ａ～２Ｄは、片側の回線に障害が発生した場合であっても、残りの一つの回線を使用して動作することができる。

【００１９】

この例において、第１、第２、第３、第４のシステムボード１Ａ、１Ｂ、１Ｃ、１Ｄは、第１のクロスバスイッチ２Ａに接続する。また、第５、第６、第７、第８のシステムボード１Ｅ、１Ｆ、１Ｇ、１Ｈは、第２のクロスバスイッチ２Ｂに接続する。また、第９、第１０、第１１、第１２のシステムボード１Ｉ、１Ｊ、１Ｋ、１Ｌは、第３のクロスバスイッチ２Ｃに接続する。第１３、第１４、第１５、第１６のシステムボード１Ｍ、１Ｎ、１Ｏ、１Ｐは、第４のクロスバスイッチ２Ｄに接続する。

【００２０】

また、図２の例において、第１のクロスバスイッチ２Ａは、バスＬ１、Ｌ２によって、第２のクロスバスイッチ２Ｂと接続する。また、第１のクロスバスイッチ２Ａは、バスＬ７、Ｌ８によって、第３のクロスバスイッチ２Ｃと接続する。また、第１のクロスバスイッチ２Ａは、バスＬ９、Ｌ１０によって、第４のクロスバスイッチ２Ｄと接続する。さらに、第２のクロスバスイッチ２Ｂは、バスＬ１１、Ｌ１２によって、第３のクロスバスイッチ２Ｃと接続し、第２のクロスバスイッチ２Ｂは、バスＬ３、Ｌ４によって、第４のクロスバスイッチ２Ｄと接続する。そして、第３のクロスバスイッチ２Ｃは、バスＬ５、Ｌ６によって、第４のクロスバスイッチ２Ｄと接続する。

【００２１】

また、各システムボード１Ａ～１Ｐも、システム制御装置（図３にて図示）を有する。情報処理システム１における各クロスバスイッチボックス２ＡＢ～２ＤＢのシステム制御装置Ｖ１～Ｖ４、及び、各システムボード１Ａ～１Ｐのシステム制御装置２２は、内部バスＬ４０によって互いに接続する。なお、図２において、情報処理システム１は、１６台のシステムボード１Ａ～１Ｐと４台のクロスバスイッチ２Ａ～２Ｄとを有するが、システムボードの台数及びクロスバスイッチの台数は１６台、４台に限定されない。続いて、各システムボード１Ａ～１Ｐの構成を説明する。

【００２２】

［システムボードの構成］

図３は、図２に示したシステムボード１Ａ～１Ｐの構成の一例を説明する図である。図３の例では、システムボード１Ａの構成を説明する。他のシステムボード１Ｂ～１Ｐの構成も、システムボード１Ａと同様である。図３に示すように、システムボード１Ａは、システムボードユニットＢ１と、サービスプロセッサボードＢ２とを有する。

【００２３】

システムボードユニットＢ１は、例えば、複数のＣＰＵ（ＣＰＵチップ）１２ａと、シ

ステムコントローラ（System Controller）１５と、Ｉ／Ｏコントローラ１６と、ＰＣＩ(Peripheral Component Interconnect）Ｅｘｐｒｅｓｓ１７と、メモリアクセスコントローラ１８と、メモリ３、１１ａと、ＭＢＣ（Maintenance Bus Controller 以下、ＭＢＣと称する）１９とを有する。メモリ３、１１ａは、例えば、ＤＲＡＭ（Dynamic Random Access Memory）である。ＭＢＣ１９は、サービスプロセッサボードＢ２との通信経路を制御する。

【００２４】

ＣＰＵ１２ａは、図１で説明したアプリケーションｐａ、ｐｂを実行する演算処理装置である。ＣＰＵ１２ａの各々は、システムコントローラ１５に接続する。システムコントローラ１５は、メモリ３、１１ａに接続されたメモリアクセスコントローラ１８に接続する。また、システムコントローラ１５は、Ｉ／Ｏコントローラ１６に接続する。Ｉ／Ｏコントローラ１６は、例えば、外部メモリ（大容量メモリ及び／又はストレージ装置）やネットワークインタフェースカード（ＮＩＣ）が接続されたＰＣＩ Ｅｘｐｒｅｓｓ１７と接続する。

【００２５】

そして、システムコントローラ１５は、ＣＰＵ１２ａとメモリアクセスコントローラ１８との間の転送制御を行う。また、システムコントローラ１５は、接続ポートａｘ、ａｙを介して、クロスバスイッチ２Ａに接続し、クロスバスイッチ２ＡとＣＰＵ１２ａとの間の転送制御、及び、クロスバスイッチ２Ａとメモリアクセスコントローラ１８との間の転送制御を行う。例えば、システムコントローラ１５は、ブリッジ回路の役割を果たす。

【００２６】

また、図１において、前述したとおり、メモリ３、１１ａの一部の領域はクロスバスイッチ２Ａを介して共有され、共有メモリ３（図１）として使用され、他の一部の領域は、ローカルメモリ１１ａとして使用される。例えば、システムコントローラ１５は、ＣＰＵ１２ａが、別のシステムボードに搭載される共有メモリ３の領域にアクセスする場合、接続ポートａｘ、ａｙを介して、クロスバスイッチ２Ａに接続する。一方、ＣＰＵ１２ａが、システムボード１Ａに搭載されるメモリ３、１１ａの領域にアクセスする場合、システムコントローラ１５は、メモリアクセスコントローラ１８にアクセスする。

【００２７】

また、サービスプロセッサボードＢ２は、システム制御装置２２とＭＢＣ（Maintenance Bus Controller 以下、ＭＢＣと称する）２１を有する。システム制御装置２２は、ノード内のハードウェアのアクセス制御、監視、電源投入、ログの採取、ユーザインターフェース制御（ユーザＩ／Ｆ）等の制御を行う。ＭＢＣ２１は、システムボードユニットＢ１との通信経路を制御する。また、ＭＢＣ２１は、ＣＰＵ１２ａやメモリ３、１１ａ、Ｉ／Ｏコントローラ１６、システム制御装置２２等のハードウェアから発生する割り込み要因を格納するレジスタｒｇを有する。また、図２で前述したとおり、システム制御装置２２は、ＬＡＮ（Local Area Network）などのネットワーク回線Ｌ４０を介して、別のノードのシステム制御装置２２、Ｖ１～Ｖ４と相互に接続する。

【００２８】

なお、図３の例では、システムボード１Ａ（１Ｂ～１Ｐ）が４台のＣＰＵ（ＣＰＵチップ）１２ａを搭載する例を示したが、システムボード１Ａが少なくとも１台のＣＰＵ１２ａを搭載する構成であっても良い。

【００２９】

続いて、図３で説明したレジスタｒｇの具体例を説明する。

【００３０】

［レジスタ］

図４は、図３のレジスタｒｇを説明する図である。図４の（Ａ）は、プロセッサのレジスタマップｒｍの一例を示す図である。また、図４（Ｂ）は、それぞれの割り込み要因の説明図である。図３で示したとおり、各ノードのサービスプロセッサボードＢ２のＭＢＣ２１は、レジスタｒｇを有する。また、レジスタｒｇは、ノードが有する複数の機能回路

（ＣＰＵ、メモリアクセスコントローラ、電源等を示す。以下、ハードウェアと称する）から発生する割り込み要因を格納する。図４の（Ａ）のレジスタマップｒｍによると、レジスタｒｇは、例えば、割り込み要因ＣＫ、ＦＥ、ＩＬ、ＥＣ、ＳＣ、ＰＭ、ＬＤ、ＩＩＯ、ＩＭを格納する。ただし、割り込み要因は、図４の例に限定されるものではない。レジスタｒｇは、それぞれの割り込み要因を、レジスタマップｒｍに対応する所定のビット位置に格納する。

【００３１】

また、図４の（Ｂ）において、割り込み要因ＣＫは、例えば、システム制御装置２２のクロック制御エラーを示す。割り込み要因ＦＥは、プロセッサにおいて発生した致命的な（ＦＡＴＡＬ）エラーを示す。また、割り込み要因ＩＬは処理対象が不正である旨のエラー、割り込み要因ＥＣはデバッグ時に使用する信号、割り込み要因ＳＣはシステム制御装置２２、Ｖ１～Ｖ４から発生したリクエスト、割り込み要因ＰＭは電源装置から発生したリクエストを示す。また、割り込み要因ＬＤはクロスバスイッチ２の二重レーンの縮退に係るエラー、割り込み要因ＩＩＯはＩ／Ｏコントローラ１６（図３）において発生するエラー、割り込み要因ＩＭはメモリアクセスコントローラ１８（図３）において発生するエラーを示す。

【００３２】

続いて、障害発生時の処理を説明する。本実施の形態例では、以下に説明する障害発生処理において、レジスタｒｇを使用する。

【００３３】

［障害発生処理］

図５は、図１～図３で述べた、本実施の形態例における情報処理システム１の一部のノードにおいて障害が発生した場合の処理の流れを説明する図である。図５において、図２、図３で示したものと同一のものは、同一の記号で示す。

【００３４】

情報処理システム１の全体の障害解析を行う場合に、複数のノードのうち、１つのノードが主体となって障害解析を行う方が効率的である。効率化のために、情報処理システム１は、１つノードのシステム制御装置をマスターのシステム制御装置に、他のノードのシステム制御装置をスレーブのシステム制御装置に設定する。または、情報処理システム１は、マスターのシステム制御装置の切り替え用として、さらに、１つノードのシステム制御装置を、マスターの代替用のシステム制御装置に設定してもよい。図５の例おいて、例えば、マスターのシステム制御装置は、クロスバスイッチ２（図１）を有する１つのノード（図２の２ＡＢ）のシステム制御装置Ｖ１である。以下、マスターのシステム制御装置Ｖ１を、マスターノード２ＡＢのシステム制御装置Ｖ１、スレーブのシステム制御装置２２を、スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４と称する。

【００３５】

前述したとおり、システム制御装置２２、Ｖ１～Ｖ４は各々、ノード内のハードウェアの状態の監視、及び、ハードウェアの制御を行う。また、システム制御装置２２、Ｖ１～Ｖ４は、ノード内の各ハードウェアから発生する割り込み要因を格納するレジスタｒｇ（図４）を有する。ハードウェアの障害の一例としては、メモリ３、１１ａのデータ破損や、プロセッサ１２ａの内部障害等が挙げられる。

【００３６】

ハードウェアの障害が発生すると割り込み信号が発生し、割り込み要因がレジスタｒｇに格納される（図示の矢印ｘ１）。システム制御装置２２、Ｖ１～Ｖ４は、レジスタｒｇを監視することによって、ハードウェアの障害の発生を検知すると、割り込み要因の発生をマスターノード２ＡＢのシステム制御装置Ｖ１に通知する（図示の矢印ｘ２、ｘ３）。続いて、マスターノード２ＡＢのシステム制御装置Ｖ１は、割り込み要因の発生の通知を受けると、各スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４に対して、ハードウェアのエラー情報を有するログ情報の送信を指示する（図示の

矢印ｘ４）。各スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４は、マスターノード２ＡＢのシステム制御装置Ｖ１からの指示に応じて、ノード内のログ情報を収集しマスターノード２ＡＢのシステム制御装置Ｖ１に送信する（図示の矢印ｘ５）。そして、各ノードにおいて取得されたログ情報が、マスターノード２ＡＢのシステム制御装置Ｖ１に収集される。

【００３７】

続いて、マスターノード２ＡＢのシステム制御装置Ｖ１は、ログ情報の解析処理を行う。例えば、システム制御装置Ｖ１は、各ノードのログ情報に基づいて、障害ノード１Ｂ、及び、障害ノード１Ｂにおける障害部品を特定する。そして、システム制御装置Ｖ１は、ログ情報の解析処理によって特定された情報に基づいて、障害に対するリアクションを行う（図示の矢印ｘ６、ｘ７）。リアクションとは、例えば、各ノードで動作するアプリケーションに対する障害ノード１Ｂが有する共有メモリ３の領域へのアクセス抑止や、障害ノード１Ｂのハードウェアの停止制御である。

【００３８】

図５で説明してきたように、一部のノードで障害が発生した場合、マスターノード２ＡＢのシステム制御装置Ｖ１は、各ノードにおいて収集されたログ情報を受信する。そして、マスターノード２ＡＢのシステム制御装置Ｖ１は、取得した各ノードのログ情報を解析することによって、障害ノード、及び、障害が発生した回路の特定処理の後、障害に対するリアクションを行う。

【００３９】

［ログ情報の解析］

図６は、図５において述べた、マスターノード２ＡＢのシステム制御装置Ｖ１におけるログ情報の解析処理（Ｓ１）の概要を説明する図である。図６において、点線で囲む工程Ｓ３、Ｓ４は、本実施の形態例において付加される処理である。

【００４０】

まず、ログ情報の解析処理（Ｓ１）を説明する。情報処理システム１は、障害が発生した場合、ノードの継続動作が可能な場合であっても、ノードの予防保守として障害の内容を特定する必要がある。また、情報処理システム１は、障害が発生しているＡＳＩＣ（Application Specific Integrated Circuit、以下、ＡＳＩＣと称する）部分を特定する必要がある。例えば、障害が発生しているＡＳＩＣ部分の特定、及び、ノードの継続動作の可否判定のために、マスターノード２ＡＢのシステム制御装置Ｖ１は、ログ情報の解析処理を行う。この実施の形態では、ＡＳＩＣは、例えば、ＣＰＵ、メモリアクセスコントローラ、Ｉ／Ｏコントローラに対応する。

【００４１】

マスターノード２ＡＢのシステム制御装置Ｖ１は、収集したログ情報に基づいてログ解析を行う（Ｓ１）。ログ情報とは、例えば、割り込み要因発生時のエラー情報を含むエラー要因情報と、エラーログ詳細情報である。エラーログ詳細情報とは、例えば、ＡＳＩＣの履歴情報やダンプ情報等である。エラーログ詳細情報はデータ量が膨大であるため、マスターノード２ＡＢのシステム制御装置Ｖ１は、ログ情報（Ｓ１）内の各解析工程Ｓ６１～Ｓ６５と平行して、エラーログ詳細情報を受信する。

【００４２】

続いて、工程Ｓ１における各解析工程を説明する。システム制御装置Ｖ１は、まず、エラー要因情報に基づいて、エラーコードの解析処理を行う（Ｓ６１）。次に、システム制御装置Ｖ１は、ノードのハードウェアそれぞれを対象として、エラー要因情報に基づいて、障害の有無の判定、及び、障害部分の特定処理を行う（Ｓ６２～Ｓ６５）。システム制御装置Ｖ１は、例えば、ＣＰＵ１２ａ、クロスバスイッチ２、メモリ３、１１ａ等を対象として、エラー要因情報に基づいて、各ハードウェアにおける障害部分の判定、及び、障害部分の詳細の判定処理を行う。工程Ｓ６２～Ｓ６５の処理により、障害ノード、及び、障害が発生している回路が特定され、他の回路が正常に動作していることが確認される。なお、図６の例において、システム制御装置Ｖ１は、ＣＰＵ１２ａ、クロスバスイッチ２

、メモリ３、１１ａを対象として解析処理を行っているが、対象となるハードウェアは、この例に限定されるものではない。

【００４３】

障害ノードの特定、及び、障害部分の特定が行われると、システム制御装置Ｖ１は、エラーログ詳細情報の収集の完了を待機して、エラーログ詳細情報の登録処理を行う（Ｓ６６）。続いて、システム制御装置Ｖ１は、エラー要因情報に基づいて、障害部分に対応するログ情報を示す代表ログの登録処理を行う（Ｓ６７）。エラーログ詳細情報、及び、代表ログは、障害の原因の分析や、障害の復旧に必要な情報である。ログ情報に基づく解析処理が完了すると、システム制御装置Ｖ１は、障害の重要度に応じて、情報処理システム１からの障害ノードの切り離し制御を行う（Ｓ２）。障害ノードの切り離し制御とは、例えば、障害ノードのハードウェアの電源停止を示す。

【００４４】

マスターノード２ＡＢのシステム制御装置Ｖ１は、障害の重要度に関わらず、ログ情報の解析処理（Ｓ１）を実行する。また、ログ情報の解析処理では、ノード内のハードウェアそれぞれを対象として、詳細に障害部分の判定処理を行うため、時間を要する。また、エラーログ詳細情報の転送処理は、エラーログ詳細情報のデータ量が膨大であるため、時間を要する。このため、ログ情報の解析処理には、数十秒～数分（３０秒～５分）程度の時間がかかる。即ち、障害の発生から障害ノードの切り離し制御まで、５分程度の時間を要する。

【００４５】

しかしながら、情報処理システム１は、障害が発生してから短時間で運用を再開することが望ましい。運用の再開処理では、正常ノードが障害ノードの処理を引き継ぐため、情報処理システム１は、障害ノードを早急に特定する必要がある。また、複数のノード間でメモリを共有する情報処理システム１では、障害の発生に起因して、共有メモリ３の破損や不整合等の二次障害が発生する恐れがある。共有メモリ３に対する二次障害を抑止するために、早急に、障害ノードのメモリへのアクセス抑止を行うことが求められる。障害発生から障害ノードのメモリへのアクセス抑止まで、例えば、１秒程度で完了することが望ましい。

【００４６】

そこで、本実施の形態例において、マスターノード２ＡＢのシステム制御装置Ｖ１は、ログ情報の解析処理（Ｓ１）の前に、ＦＮＬ（Fail Node List、以下、ＦＮＬと称する）解析処理（Ｓ３）を行って障害ノードを特定し、障害ノードのメモリへのアクセス抑止を行う（Ｓ４）。

【００４７】

本実施の形態例のマスターノード２ＡＢのシステム制御装置Ｖ１は、他のノードの割り込み要因の発生に応じてレジスタｒｇの割り込み要因を受信し、割り込み要因のうち、障害として検出すべき割り込み要因を抽出する。そして、マスターノード２ＡＢのシステム制御装置Ｖ１は、抽出結果に応じて障害ノードを特定し、障害ノードのメモリへのアクセスを抑止後、他のノードから受信したログ情報に基づいて障害ノードの切り離し制御を行う。

【００４８】

具体的に、マスターノード２ＡＢのシステム制御装置Ｖ１は、ＦＮＬ解析処理（Ｓ３）として、まず、発生中の割り込み要因を各ノードから取得する（Ｓ５１）。続いて、システム制御装置Ｖ１は、取得した割り込み要因のうち、障害として検出すべき割り込み要因を抽出する（Ｓ５２）。次に、システム制御装置Ｖ１は、抽出した割り込み要因のうち、波及先の割り込み要因をＦＮＬ解析の対象から除外する（Ｓ５３）。即ち、システム制御装置Ｖ１は、抽出した割り込み要因のうち、別の割り込み要因に起因して発生した割り込み要因を、ＦＮＬ解析処理の対象外とする。

【００４９】

続いて、システム制御装置Ｖ１は、複数の割り込み要因が抽出された場合、各割り込み

要因の優先度を判定する（Ｓ５４）次に、システム制御装置Ｖ１は、優先度の高い順に割り込み要因を選択し、割り込み要因に対応する障害ノードを特定する（Ｓ５５）。次に、システム制御装置Ｖ１は、障害ノードの他ノードからのメモリへのアクセス抑止処理を行う（Ｓ５６）。即ち、システム制御装置Ｖ１は、障害ノードが有する共有メモリ３の領域に対するアクセスを抑止する。各工程の詳細については、後述する。続いて、システム制御装置Ｖ１は、ログ情報の解析処理を実行し（Ｓ１）、障害ノードの情報処理システム１からの切り離し制御を行う（Ｓ２）。

【００５０】

図６で説明してきたように、本実施の形態例において、システム制御装置Ｖ１は、ＦＮＬ解析処理（Ｓ３）として、ログ情報の代わりに割り込み要因に基づいて、障害ノードを特定し、障害ノードのメモリに対する他ノードからのアクセス抑止処理を行う（図６のＳ５６）。障害ノードのメモリに対するアクセスを抑止することによって、システム制御装置Ｖ１は、共有メモリ３の二次障害を早急に抑止し、障害発生時の障害ノードによる他ノードへの影響を低減する。

【００５１】

そして、システム制御装置Ｖ１は、障害ノードのメモリへのアクセス抑止後、ログ情報の解析処理（Ｓ１）を行って、障害が発生しているＡＳＩＣ部分を特定し、ノードの継続動作の可否を判定する。そして、システム制御装置Ｖ１は、ログ情報の解析処理の結果に基づいて、障害ノードの情報処理システム１からの切り離し制御（Ｓ２）を行う。

【００５２】

図７は、図６のログ情報の解析処理（図６のＳ１）、及び、ＦＮＬ解析処理（Ｓ３）に要する時間を例示する図である。図７の（Ａ）は、ログ情報の解析処理（Ｓ１）から障害ノードの切り離し制御（Ｓ２）までの時間を示す図であって、図７の（Ｂ）は、ＦＮＬ解析処理（Ｓ３）から障害ノードのメモリへのアクセス抑止処理（Ｓ４）までの時間を示す図である。

【００５３】

図７の（Ａ）では、ログ情報の解析処理（Ｓ１）の後、障害ノードの切り離し制御（Ｓ２）が行われる。前述したとおり、ログ情報の解析処理（Ｓ１）は、ハードウェアそれぞれに対するログ情報の解析処理やエラーログ詳細情報の転送処理に伴って、時間を要する。図７の（Ａ）によると、障害発生から障害ノードの切り離し制御まで、期間ｔ１に示す時間を要する。

【００５４】

一方、図７の（Ｂ）において、ＦＮＬ解析処理（Ｓ３）では、システム制御装置Ｖ１は、障害として検出すべき割り込み要因に基づいて障害ノードを特定する。また、ＦＮＬ解析処理（Ｓ３）では、エラーログ詳細情報の転送が不要であり、割り込み要因（３２ビット程度）のデータ量は小さい。したがって、システム制御装置Ｖ１は、障害ノードを早急に特定することが可能になるため、図７の（Ｂ）によると、障害発生から障害ノードのメモリへのアクセス抑止処理までの時間ｔ２は、時間ｔ１に対して大幅に短縮される。

【００５５】

ここで、本実施の形態例におけるマスターノード２ＡＢのシステム制御装置Ｖ１、及び、スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４のソフトウェアモジュール図を説明する。

【００５６】

［ソフトウェアモジュール図］

図８は、本実施の形態における情報処理システムの各ノードのソフトウェアモジュール図である。図８は、マスターノード２ＡＢのシステム制御装置Ｖ１、及び、スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４のブロック図を有する。初めに、スレーブノード１Ａ～１Ｐ、２ＢＢ～２ＤＢのシステム制御装置２２、Ｖ２～Ｖ４のブロックを説明する。ここでは、スレーブノード１Ａのシステム制御装置２２について説明する。

【００５７】
図８において、スレーブノード１Ａのシステム制御装置２２は、例えば、ＦＮＬ（Fail Node List）ドライバ５４、ＦＮＬ（Fail Node List）部５０、ハード内制御部６１、ＲＡＳ（Reliability Availability Serviceability、以下、ＲＡＳと称する）６２、ＸＳＣＦ（eXtended System Control Facility、以下、ＸＳＣＦと称する）コマンド部６３、ハイパーバイザ６４を有する。また、ＦＮＬ部５０は、例えば、ＦＮＬ（Fail Node List）制御部５１、ＦＮＬ（Fail Node List）更新依頼受信制御部５２、ＦＮＬ（Fail Node List）更新部５３を有する。

【００５８】
ハード内制御部６１は、例えば、電源やプロセッサ（図８では、ＣＰＵと記す）やクロスバスイッチ（図８では、ＸＢと記す）等のハードウェアに対するアクセス処理を行うＨＡＰ（Hardware Access Program、以下、ＨＡＰと称する）６５を有する。そして、ハード内制御部６１は、当該ハードウェアに対するアクセス処理における割り込み要因の発生を検知し、ＦＮＬ部５０のＦＮＬ更新依頼受信制御部５２に通知する。また、ＲＡＳ６２は、ＡＳＩＣにおける割り込み要因の発生を検知し、ＦＮＬ更新依頼受信制御部５２に通知する。また、ＸＳＣＦコマンド部４３は、例えば、ハイパーバイザにおける割り込み要因の発生を検知し、ＦＮＬ更新依頼受信制御部５２に通知する。

【００５９】
また、ＦＮＬ更新依頼受信制御部５２は、各部からの割り込み要因の発生の通知を取得して、ＦＮＬ制御部５１に出力する。そして、ＦＮＬ制御部５１は、ＦＮＬドライバ５４を介して、割り込み要因の発生をマスターノード２ＡＢのシステム制御装置Ｖ１のＦＮＬ制御部３２に通知する。また、ＦＮＬ制御部５１は、マスターノード２ＡＢのシステム制御装置Ｖ１からの割り込み要因の収集依頼に応答して、発生している割り込み要因を収集しＦＮＬドライバ５４を介してマスターノード２ＡＢのシステム制御装置Ｖ１に送信する。ＦＮＬ更新部５３は、マスターノード２ＡＢのシステム制御装置Ｖ１からのＦＮＬ更新指示に基づいて、ＦＮＬ（Fail Node List、図８には図示せず）を更新する。ＦＮＬとは、メモリを共有するノードそれぞれのアクセス処理の可否を管理するリストである。情報処理システム１における各ノードは、ＦＮＬに基づいて、アクセス抑止対象のノードを検知する。

【００６０】
また、図８において、マスターノード２ＡＢのシステム制御装置Ｖ１は、例えば、ＦＮＬドライバ３５、ＦＮＬ部３０、ＸＳＣＦコマンド部４３、ハード内制御部４１、ＲＡＳ４２、ＦＮＤＢ３６（Fail Node DB、以下、ＦＮＤＢと称する）を有する。また、ＦＮＬ部３０は、例えば、ＦＮＬ解析部３１、ＦＮＬ制御部３２、ＦＮＬ更新部３３、ＦＮＬ更新依頼受信制御部３４を有する。ハード内制御部４１、ＲＡＳ４２、ＸＳＣＦコマンド部４３、ＦＮＬ更新依頼受信制御部３４の処理は、スレーブノード１Ａのシステム制御装置２２と同様である。

【００６１】
ＦＮＬ部３０のＦＮＬ制御部３２は、スレーブノード１Ａのシステム制御装置２２から割り込み要因発生の通知を受信すると、ＦＮＬドライバ３５を介して、各スレーブノード１Ａのシステム制御装置２２に対して、発生中の割り込み要因の収集を指示する。ＦＮＬ解析部３１は、各ノードのシステム制御装置２２から収集した割り込み要因に基づいて、ＦＮＤＢ３６を参照し障害ノードを特定する。ＦＮＤＢ３６は、ＦＮＬ解析における解析論理の定義を有するファイルである。そして、ＦＮＬ解析部３１は、特定した障害ノードの情報に基づいて、各ノードのＦＮＬ更新部３３にＦＮＬの更新を指示する。

【００６２】
図９は、図８において説明したＦＮＬ（Fail Node List）４０の一例を示す図である。図９例において、情報処理システム１は、例えば、図２で示したように、１６個のノードＳＢ００～ＳＢ１５を有し、各ノードはメモリを共有する。そこで、図９のＦＮＬ４０は、１６個のノードそれぞれに対するアクセス処理の可否を管理する値を有する。例えば、

値「０」の場合、対象ノードの共有メモリ３に対するアクセスが許可されることを示す。一方、値「１」の場合、対象ノードの共有メモリ３に対するアクセスが抑止されることを示す。

【００６３】

続いて、割り込み要因が発生した後、図９において説明したＦＮＬ４０が更新されるまでの処理の流れを、図８において説明したソフトウェアモジュールに対応して、時系列に説明する。

【００６４】

［ソフトウェアモジュールの処理の流れ］

図１０は、割り込み要因が発生しＦＮＬが更新される間のマスターノード２ＡＢのシステム制御装置Ｖ１、及び、スレーブノード１Ａのシステム制御装置２２における、処理の流れを時系列に説明する図である。図１０において、図８で示したものと同一のものは同一の記号で示してある。

【００６５】

図１０の例において、例えば、スレーブノードの一部のＡＳＩＣにおいて障害が発生する。障害の発生により割り込み信号が発生し、障害が発生したハードウェアに対応する割り込み要因がレジスタｒｇに登録される。スレーブノード１Ａのシステム制御装置２２は、障害の発生を検知すると（図示の矢印ｇ１）、障害が発生したことをマスターノード２ＡＢのシステム制御装置Ｖ１におけるハード内制御部４１に通知する（図示の矢印ｇ２）。

【００６６】

システム制御装置Ｖ１におけるハード内制御部４１は、障害の発生の通知を受けて、各スレーブノード１Ａのシステム制御装置２２に対して割り込み要因の収集を指示する（図示の矢印ｇ３）。マスターノード２ＡＢのシステム制御装置Ｖ１の通知に応答して、各スレーブノード１Ａのシステム制御装置２２は、発生している割り込み要因を、ＦＮＬ部５０を介して（図示の矢印ｇ４）、システム制御装置Ｖ１のＦＮＬ部３０に送信する（図示の矢印ｇ５）。

【００６７】

この結果、各スレーブノード１Ａのシステム制御装置２２の割り込み要因が収集される。割り込み要因のデータ量は小さい。このため、システム制御装置Ｖ１は、短時間で、各スレーブノード１Ａのシステム制御装置２２の割り込み要因を取得することができる。また、各システム制御装置２２、Ｖ１～Ｖ４間が高速通信を介して接続される場合、システム制御装置Ｖ１は、さらに、高速に、各システム制御装置２２、Ｖ２～Ｖ４の割り込み要因を取得することができる。

【００６８】

マスターノード２ＡＢのシステム制御装置Ｖ１におけるＦＮＬ部３０は、全てのスレーブノード１Ａにおける割り込み要因を収集すると、ＦＮＬ解析部３１に解析処理を指示する（図示の矢印ｇ６）。そして、ＦＮＬ解析部３１は、収集した割り込み要因に基づいて障害ノードを特定し、ＦＮＬ部３０に出力する（図示の矢印ｇ７）。続いて、ＦＮＬ部３０は、障害ノードの情報に基づいて、各スレーブノード１Ａのシステム制御装置のＦＮＬ部５０にＦＮＬ４０（図９）の更新を指示する（図示の矢印ｇ８）。ＦＮＬ４０の更新指示を受信すると、スレーブノード１ＡのＦＮＬ部５０は、ＦＮＬ更新部５３にＦＮＬ４０の更新を実行させる（ｇ９、ｇ１０）。

【００６９】

図１０のフローチャート図に示した処理の流れに基づいて、ＦＮＬ解析処理、及び、ＦＮＬの更新処理が行われる。続いて、各処理の詳細をフローチャート図に基づいて説明する。

【００７０】

［ＦＮＬ解析処理、ＦＮＬ更新処理］

図１１は、図８の本実施の形態例におけるＦＮＬ解析部３１の処理、及び、ＦＮＬ更新

部３３の処理を説明するフローチャート図である。初めに、例えば、マスターノード２ＡＢのシステム制御装置Ｖ１におけるＦＮＬ解析部３１は、電源障害が発生しているか否かを判定する（Ｓ２１）。電源障害が発生している場合、電源障害の対応が優先されるため、ＦＮＬ解析部３１は処理を終了する。

【００７１】

一方、電源障害が発生していない場合（Ｓ２１のＮＯ）、ＦＮＬ解析部３１は、割り込み要因を取得する（Ｓ２２）。前述したとおり、ＦＮＬ解析部３１は、割り込み要因が発生したノードからの通知に応答して、各ノードから割り込み要因を取得する。続いて、ＦＮＬ解析部３１は、収集された割り込み要因から、障害として検出すべき割り込み要因を抽出する（Ｓ２３）。例えば、ＦＮＬ解析部３１は、例えば、割り込み要因のうち、ノードの停止が必要となる障害に対応する割り込み要因を抽出する。即ち、ＦＮＬ解析部３１は、ノードが継続して動作可能な割り込み要因を抽出の対処としない。

【００７２】

本実施の形態例において、マスターノード２ＡＢのシステム制御装置Ｖ１は、障害として検出すべき割り込み要因として、例えば、図４に例示した割り込み要因のうち、割り込み要因ＣＫ、ＦＥを抽出する（Ｓ２３）。割り込み要因ＣＫ、ＦＥは、ＣＰＵが停止する障害要因であって、割り込み要因ＣＫ、ＦＥ以外の割り込み要因については機能の一部が縮退する故障要因であるためである。ただし、この例に限定されるものではなく、システム制御装置Ｖ１は、別の割り込み要因を、障害として検出すべき割り込み要因としてもよい。

【００７３】

なお、図４のレジスタマップｒｍは、プロセッサのレジスタマップである。情報処理システム１では、プロセッサのレジスタマップの他に、クロスバスイッチ用のレジスタマップやＭＢＣ用のレジスタマップが存在する。また、クロスバスイッチ用のレジスタに格納される割り込み要因については、例えば、複数の割り込み要因のうち、内部障害、及び、ポート障害に対応する割り込み要因が抽出対象となる。また、ＭＢＣ用のレジスタに格納される割り込み要因については、例えば、複数の割り込み要因のうち、割り込み要因ＦＥが抽出対象となる。

【００７４】

次に、ＦＮＬ解析部３１は、波及先の割り込み要因を抑止する（Ｓ２４）。ここで、割り込み要因は、波及元の割り込み要因と、波及元の割り込み要因に基づいて誘発された波及先の割り込み要因とに区分される。ＦＮＬ解析部３１は、波及先の割り込み要因を除外して、波及元の割り込み要因のみに絞り込む。

【００７５】

具体的に、例えば、あるノードのプロセッサにおいて障害が発生した場合、同一ノード内のクロスバスイッチの接続部や、プロセッサの他の部分に障害が波及することがある。この場合、プロセッサにおいて発生した障害に対応する割り込み要因が上位の割り込み要因、クロスバスイッチの接続部やプロセッサの他の部分において発生した障害に対応する割り込み要因が下位の割り込み要因に相当する。即ち、上位の割り込み要因が波及元の割り込み要因に該当し、下位の割り込み要因が波及先の割り込み要因に該当する。

【００７６】

ＦＮＬ解析部３１は、波及先の割り込み要因の抑止処理（Ｓ２４）によって、波及先の割り込み要因をＦＮＬ解析の対象から除外するため、波及元の割り込み要因に対応するノードのみを障害ノードとして特定する。即ち、ＦＮＬ解析部３１は、波及先の割り込み要因に対応するノードを障害ノードして特定することを回避し、真に障害が発生するノードのみを、障害ノードとして特定する。

【００７７】

続いて、ＦＮＬ解析部３１は、抽出した割り込み要因それぞれの優先度を取得する（Ｓ２５）。そして、ＦＮＬ解析部３１は、優先度の高い割り込み要因から順に、当該割り込み要因に対応して、障害ノードを特定すると共に、障害ノードに対する制御内容を取得す

る（Ｓ２６）。そして、ＦＮＬ更新部３３、５３は、取得した制御内容に基づいてＦＮＬ４０を更新し、共有メモリ３上の障害ノードの領域への、他ノードからのアクセスを抑止する（Ｓ２７）。続いて、ＦＮＬ解析部３１、及び、ＦＮＬ更新部３３、５３は、次に優先度の高い割り込み要因を対象として、工程Ｓ２６、Ｓ２７の処理を行う。そして、ＦＮＬ解析部３１、及び、ＦＮＬ更新部３３、５３は、抽出した全ての割り込み要因を対象として工程Ｓ２６、Ｓ２７の処理を行うと、ＦＮＬの解析処理、及び、ＦＮＬの更新処理を終了する。

【００７８】

続いて、図１１のフローチャート図における波及先の割り込み要因の抑止処理（工程Ｓ２４）の詳細を説明する。

【００７９】

［波及先割り込み要因の抑止（図１１の工程Ｓ２４）］

図１２は、波及先の割り込み要因の抑止処理を説明するフローチャート図である。まず、ＦＮＬ解析部３１は、ＦＮＤＢ３６を参照し、抽出した割り込み要因（この例では、ＣＫ、ＦＥ）が、波及元の割り込み要因であるか否かを判定する（Ｓ１１）。ＦＮＤＢ３６については、次の図１３に基づいて説明する。そして、抽出した割り込み要因が、下位の割り込み要因（波及先の割り込み要因）ではない場合（Ｓ１２のＮＯ）、即ち、波及元の割り込み要因である場合、ＦＮＬ解析部３１は波及先の割り込み要因の消し込み処理を終了する。

【００８０】

一方、抽出した割り込み要因が下位の割り込み要因である場合（Ｓ１２のＹＥＳ）、ＦＮＬ解析部３１は、各ノードのレジスタｒｇを参照し、当該下位の割り込み要因に対応する上位の割り込み要因を示す、波及元の割り込み要因が発生しているか否かを判定する（Ｓ１３）。波及元の割り込み要因が発生している場合（Ｓ１４のＹＥＳ）、ＦＮＬ解析部３１は、下位の割り込み要因を抑止し、ＦＮＬ解析処理の対象外とする（Ｓ１５）。

【００８１】

一方、波及元の割り込み要因が発生していない場合（Ｓ１４のＮＯ）、ＦＮＬ解析部３１は波及先の割り込み要因の抑止処理を終了する。つまり、例えば、発生した割り込み要因が波及先の割り込み要因である場合であっても、波及元の割り込み要因が発生していない場合、ＦＮＬ解析部３１は、波及先の割り込み要因を抑止しない。

【００８２】

図４、図１２で述べてきたとおり、本実施の形態例では、ＦＮＬ解析部３１は、割り込み要因のうち、ノードの停止が必要となる障害に対応する割り込み要因に限定して、ＦＮＬ解析を行う（図１１のＳ２３）。また、ＦＮＬ解析部３１は、さらに、波及元となる割り込み要因に限定して、ＦＮＬ解析を行う（図１１のＳ２４）。したがって、ＦＮＬ解析部３１は、ノードの停止が必要となる障害に対応する割り込み要因であって、真に障害が発生しているノードにおける割り込み要因を抽出することができる。即ち、ＦＮＬ解析部３１は、ノードの停止が必要となる最小限の障害ノードを、効率的に特定することができる。

【００８３】

ここで、ＦＮＬ解析部３１が、ＦＮＬ解析処理において参照するＦＮＤＢ３６の具体例を説明する。ＦＮＤＢ３６は、ＦＮＬ解析における解析論理の定義を有する。

【００８４】

［ＦＮＤＢの具体例］

図１３は、ＦＮＤＢ３６の具体例を示す図である。図１３の（Ａ）は、ＦＮＬ解析における解析論理を定義する定義テーブルｔｂ１を示す図である。図１３の（Ｂ）は、定義テーブルｔｂ１に記述される各エントリの一部を説明する図である。定義テーブルｔｂ１は、例えば、共通定義フレームとデータ定義ブロックとを有する。共通定義フレームは、定義テーブルｔｂ１の版数や定義開始の宣言を有する。

【００８５】

図１３の（Ａ）によると、定義テーブルｔｂ１には、例えば、割り込み要因に対応して、優先度（ｐｒｉｏ）、アクション番号（ａｃｔ）、エントリ抑止条件（ｅｎｔ_ｄｉｓ）等の定義を有する。また、図１３の（Ｂ）によると、優先度（ｐｒｉｏ）は、割り込み要因の優先度を示す。例えば、優先度は数値で示される。また、アクション番号（ａｃｔ）は、割り込み要因に対応する、障害ノード及び共有メモリへのアクセス抑止処理の種別を示す。アクセス抑止処理の種別の詳細については、別の図１４に基づいて説明する。

【００８６】

そして、エントリ抑止条件（ｅｎｔ_ｄｉｓ）は、割り込み要因に対応して、当該割り込み要因に対して論理的に上位の割り込み要因を示す。即ち、エントリ抑止条件（ｅｎｔ_ｄｉｓ）は、割り込み要因に対応する、波及元の割り込み要因を示す。エントリ抑止条件（ｅｎｔ_ｄｉｓ）がブランクの場合、割り込み要因に対応する、波及元の割り込み要因が存在しないことを示す。

【００８７】

ＦＮＬ解析部３１は、例えば、図１３の定義テーブルｔｂ１の記述ｃｄ１を参照して、波及先の割り込み要因の抑止処理（図１１のＳ２４）、優先度の取得処理（図１１のＳ２５）、障害ノードの特定、及び、アクセス抑止処理の取得処理（図１１のＳ２６）を行う。例えば、ＦＮＬ解析部３１は、定義テーブルｔｂ１内のａｄｒｓ（割り込み要因番号）列ｃｄ２を参照し、ａｄｒｓの値が、割り込み要因に対応する割り込み要因番号と一致する行を探索する。

【００８８】

例えば、レジスタｒｇに基づいて収集した割り込み要因が、割り込み要因ＣＫである場合を例に挙げる。割り込み要因ＣＫは、図４のレジスタマップｒｍによると、１ビット目の位置に位置することから、割り込み要因ＣＫに対応する割り込み要因番号は値「０ｘ０００００００１」である。そこで、ＦＮＬ解析部３１は、割り込み要因ＣＫに対応して、定義テーブルｔｂ１における２行目の定義情報を検出する。なお、割り込み要因ＦＥに対応する割り込み要因は値「０ｘ０００００００３」である。そこで、ＦＮＬ解析部３１は、割り込み要因ＦＥに対応して、定義テーブルｔｂ１における１行目の定義情報を検出する。

【００８９】

続いて、ＦＮＬ解析部３１は、検出した２行目の定義情報のうち、エントリ抑止条件（ｅｎｔ_ｄｉｓ）に対応する項目ｃｄ５を参照する。図１３の例において、２行目の定義情報におけるエントリ抑止条件（ｅｎｔ_ｄｉｓ）は、ブランクである。したがって、ＦＮＬ解析部３１は、割り込み要因ＣＫを、波及元の割り込み要因として判定する（図１１のＳ２４）。また、ＦＮＬ解析部３１は、２行目の定義情報における優先度（ｐｒｉｏ）の項目ｃｄ３に基づいて、優先度「０ｘ０１」（図１１のＳ２５）を取得すると共に、アクション番号（ａｃｔ）の項目ｃｄ４に基づいてアクション番号「０ｘ０１」を取得する。なお、図１３の例における優先度は、値が小さいほど高い。

【００９０】

一方、ＦＮＬ解析部３１は、割り込み要因ＦＥに対応して、１行目の定義情報を検出する。図１３の例において、定義テーブルｔｂ１は、１行目の定義情報におけるエントリ抑止条件（ｅｎｔ_ｄｉｓ）として、定義“/ＸＢＢＯＸ/ＸＢＵＸ/ＧＸＢ/ＦＮ_ＸＢ_ＳＮＤ”の記述を有する（ｃｄ５）。それぞれの定義ＸＢＢＯＸ、ＸＢＵＸ、ＧＸＢ、ＦＮ_ＸＢ_ＳＮＤは、割り込み要因を示し、割り込み要因ＦＥの波及元の割り込み要因に該当する。定義ＸＢＢＯＸは、例えば、クロスバボックスにおける割り込み要因を、定義ＦＮ_ＸＢ_ＳＮＤは、例えば、クロスバスイッチの送信部における割り込み要因を示す。したがって、ＦＮＬ解析部３１は、割り込み要因ＦＥを波及先の割り込み要因として判定し、ＦＮＬ解析の対象から除外する。

【００９１】

具体例に基づいて説明する。例えば、複数のノードＳＢ００～ＳＢ０３が、クロスバスイッチ２を備えるノードＸＢ００と接続する情報処理システムを例示する。具体例では、

例えば、ノードＳＢ０２においてクロック制御エラーが発生すると共に、クロスバスイッチ２を備えるノードＸＢ００において、ポート障害が発生する。

【００９２】

いずれかの割り込み要因を検知すると、マスターノード２ＡＢのシステム制御装置Ｖ１は、各ノードにおいて発生する割り込み要因を収集する（Ｓ２２）。そして、システム制御装置Ｖ１は、ノードＳＢ０２において発生したクロック制御エラーに対応する割り込み要因ＣＫと、ノードＸＢ００において発生したポート障害に対応する対応する割り込み要因を取得する。具体例において、割り込み要因ＣＫ、及び、ポート障害に対応する割り込み要因は、抽出対象の割り込み要因である（Ｓ２３）。

【００９３】

続いて、ＦＮＬ解析部３１は、図１３のＦＮＤＢ３６を参照して（ｃｄ５）、各割り込み要因が波及元の割り込み要因であるか否かを判定する。図１３において前述したとおり、割り込み要因ＣＫは、波及元の割り込み要因である。また、図示していないが、具体例において、ポート障害に対応する割り込み要因は、波及元の割り込み要因である。このため、ＦＮＬ解析部３１は、ポート障害に対応する割り込み要因を抑止しない（Ｓ２４）。次に、ＦＮＬ解析部３１は、ＦＮＤＢ３６を参照して（ｃｄ３）、各割り込み要因に対応する優先度を取得する（Ｓ２５）。図１３において前述したとおり、割り込み要因ＣＫの優先度（ｐｒｉｏ）は、優先度「０ｘ０１」である。また、具体例において、図示していないが、ポート障害に対応する割り込み要因の優先度は、優先度「０ｘ０５」である。したがって、ＦＮＬ解析部３１は、割り込み要因ＣＫを、ポート障害に対応する割り込み要因よりも優先する。

【００９４】

続いて、図１３の定義テーブルｔｂ１のアクション番号（ａｃｔ）を説明する。アクション番号（ａｃｔ）は、割り込み要因に対応する、障害ノード及び共有メモリへのアクセス抑止処理の種別を示す。図１３の定義テーブルｔｂ１の記述ｃｄ４によると、割り込み要因ＣＫのアクション番号（ａｃｔ）は、「０ｘ０１」である。また、図示していないが、ポート障害に対応する割り込み要因のアクション番号（ａｃｔ）は、例えば、「０ｘ１２」である。次の図１４に基づいて、アクション番号（ａｃｔ）に対応する制御情報について説明する。

【００９５】

図１４は、アクション番号（ａｃｔ）を有する定義テーブルｔｂ２の具体例を示す図である。図１４の（Ａ）は、アクション番号（ａｃｔ）に対応して制御情報（ｒｕｌｅ）の記述を有する定義テーブルｔｂ２を示す図であって、図１４の（Ｂ）は、定義テーブルｔｂ２に記述される制御情報（ｒｕｌｅ）の各エントリの一部を説明する図である。ＦＮＤＢ３６は、例えば、図１３の定義テーブルｔｂ１に加えて、図１４に示す定義テーブルｔｂ２を有する。定義テーブルｔｂ２は、例えば、定義テーブルｔｂ２の版数や定義開始の宣言を有する共通定義フレームと、データ定義ブロックとを有する。

【００９６】

図１４の定義テーブルｔｂ２のデータ定義ブロックは、アクション番号（ａｃｔ）に対応して、障害ノードに対する制御情報（ｒｕｌｅ）の記述ｃｄ６を有する。例えば、定義テーブルｔｂ２は、アクション番号（ａｃｔ）「０ｘ０１」に対応して、制御情報（ｒｕｌｅ）として、エントリＦＮＬ_ＵＰＤＡＴＥを有する。また、定義テーブルｔｂ２は、アクション番号（ａｃｔ）「０ｘ０２」に対応して、制御情報（ｒｕｌｅ）として、エントリＦＮＬ_ＵＰＤＡＴＥ_ＤＥＳＴを有する。同様にして、定義テーブルｔｂ２は、アクション番号（ａｃｔ）「０ｘ１１」に対応して、制御情報（ｒｕｌｅ）として、エントリＧＣＳＭ_ＤＥＧＲＡＤＥを、アクション番号（ａｃｔ）「０ｘ１２」に対応して、制御情報（ｒｕｌｅ）として、エントリＧＣＳＭ_ＤＥＧＲＡＤＥ_ＤＥＳＴを有する。

【００９７】

図１４の（Ｂ）によると、エントリＦＮＬ_ＵＰＤＡＴＥは、割り込み要因が検出されたノードを障害ノードとして特定し、当該障害ノードを停止対象のノードとしてメモリア

クセスの制御を行うことを示す。この場合、マスターノード２ＡＢのシステム制御装置Ｖ１は、例えば、割り込み要因が検出されたノード（障害ノード）の共有メモリ３内の領域に対する、他のノードからのアクセスを抑止する。アクセスが抑止されることにより、障害ノードのメモリが共有メモリから切り離され、情報処理システムの継続稼動が可能になる。また、図１４の（Ｂ）によると、エントリＦＮＬ_ＵＰＤＡＴＥ_ＤＥＳＴは、割り込み要因が検出されたノードに接続されたノードを障害ノードとして特定し、当該障害ノードを停止対象のノードとしてメモリアクセスの制御を行うことを示す。

【００９８】

さらに、エントリＧＣＳＭ_ＤＥＧＲＡＤＥは、割り込み要因が検出されたノードを障害ノードとして特定し、当該障害ノードを機能縮退対象のノードとしてメモリアクセスの制御を行うことを示す。さらに、エントリＧＣＳＭ_ＤＥＧＲＡＤＥ_ＤＥＳＴは、割り込み要因が検出されたノードに接続されたノードを障害ノードとして特定し、当該障害ノードを機能縮退対象のノードとしてメモリアクセスの制御を行うことを示す。障害ノードの機能縮退とは、例えば、障害ノードがクロスバスイッチ２を備えるノードである場合に、図２で説明したクロスバスイッチの二重の回線を一回線に縮退させる制御を示す。

【００９９】

具体例において、前述したとおり、割り込み要因ＣＫのアクション番号（ａｃｔ）は、値「０ｘ０１」であって、ポート障害に対応する割り込み要因のアクション番号（ａｃｔ）は、例えば、値「０ｘ１２」である。したがって、ＦＮＬ解析部３１は、割り込み要因ＣＫに対応して割り込み要因が発生したノードＳＢ０２を障害ノードとして特定する。そして、ＦＮＬ更新部３３、５３は、ノードＳＢ０２に係るメモリアクセスの制御（ＦＮＬ_ＵＰＤＡＴＥ）を行う。また、ＦＮＬ解析部３１は、ポート障害に対応する割り込み要因が発生したノードＸＢ００に接続されるノードＳＢ００～ＳＢ０３を障害ノードとして特定する。そして、ＦＮＬ更新部３３、５３は、ノードＳＢ００～ＳＢ００に係る機能縮退制御（ＧＣＳＭ_ＤＥＧＲＡＤＥ_ＤＥＳＴ）を行う。

【０１００】

ただし、具体例において、割り込み要因ＣＫは、ポート障害に対応する割り込み要因よりも優先される。そこで、ＦＮＬ更新部３３、５３は、まず、割り込み要因ＣＫに対応する障害ノードのメモリへのアクセス抑止処理を行う（Ｓ２６、Ｓ２７）。例えば、それぞれのノードのＦＮＬ更新部３３、５３はＦＮＬ４０を更新し、ノードＳＢ０２のメモリに対する他のノードからのアクセスを抑止する。

【０１０１】

続いて、ＦＮＬ更新部３３、５３は、ポート障害に対応する割り込み要因に対応して、ノードＳＢ００～ＳＢ０３のメモリへのアクセス抑止処理を行う（Ｓ２６、Ｓ２７）。例えば、それぞれのノードのＦＮＬ更新部３３、５３はＦＮＬ４０を更新し、ノードＸＢ００からノードＳＢ００～ＳＢ０３に対するアクセス処理における二重の回線を片側の回線に縮退させる。回線が縮退されたことにより、ノードＳＢ００～ＳＢ０３の共有メモリに対するアクセス経路が減少する。

【０１０２】

図１５は、具体例におけるメモリの抑止範囲を説明する図である。具体例によると、ノードＳＢ０２においてクロック制御エラーが発生した場合、他ノードによる、ノードＳＢ０２の共有メモリ３に対するアクセスが抑止される（ａｃ１）。一方、ノードＸＢ００においてポート障害が発生した場合、ノードＸＢ００とノードＳＢ００～ＳＢ０３との間の回線が片側に縮退される（ａｃ２）。即ち、図１５の例において、回線ｎ１、ｎ３、ｎ５、ｎ７が使用できない状態となる。なお、例えば、回線ｎ１、ｎ３、ｎ５、ｎ７が既に停止されている状態で、さらに、ポート障害が発生した場合、全ての回線ｎ１～ｎ８が使用不可状態となり、ノードＳＢ００～ＳＢ０３の共有メモリ３に対するアクセスが行えなくなる。

【０１０３】

図１５に示すように、割り込み要因ＣＫに対応するアクセスの抑止範囲は、クロスバス

イッチ２のポート障害によるアクセスの抑止範囲より狭い。図１５の例では、アクセスの抑止範囲のより広いポート障害に対応する割り込み要因の優先度が低く設定されることにより、ポート障害によるアクセス抑止処理は、割り込み要因ＣＫによるアクセス抑止処理より後から行われる。アクセスの抑止範囲の広い割り込みがより後から行われることによって、情報処理システム１の性能がより長時間維持される。図１５の例のように、例えば、割り込み要因の優先度は、システム制御装置１の性能をより高性能に維持するために、抑止範囲がより小さい割り込み要因ほど、より高い優先度が設定される。

【０１０４】

以上のように、本実施の形態例における情報処理システムは、ノードの各々は、複数の機能回路と機能回路を制御する制御装置と、複数の機能回路から発生する割り込み要因を格納するレジスタとを有する。また、情報処理システムにおける複数のノードのうちの１のノードの制御装置は、他のノードの割り込み要因の発生に応じてレジスタの割り込み要因を受信し、割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定する。そして、制御装置は、障害ノードのメモリへのアクセスを抑止後、他のノードから受信したログ情報に基づいて障害ノードの切り離し制御を行う。

【０１０５】

本実施の形態例における情報処理システムは、割り込み要因に基づくことにより障害ノードを高速に特定することができる。また、本実施の形態例における情報処理システムは、複数の割り込み要因のうち、ノードの停止が必要となる、障害として検出すべき割り込み要因を対象として、障害ノードを特定するため、より効率的に、障害ノードを特定することができる。

【０１０６】

また、本実施の形態例における情報処理システムは、高速に、障害ノードを特定することができるため、障害ノードのメモリへのアクセスを早急に抑止することができ、共有メモリへの二次障害を回避することができる。即ち、情報処理システムは、障害発生時に障害ノードによる他ノードの影響を早急に低減することができる。また、情報処理システムは、障害ノードを高速に特定できるため、障害の発生時における、障害ノードから正常ノードへの運用の切り替えにかかるオーバヘッドを低減できる。

【０１０７】

また、本実施の形態例における情報処理システムにおいて、１のノードの制御装置は、障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定し、発生していない場合に、割り込み要因に対応するノードを障害ノードとして特定し、発生している場合に、波及元となる割り込み要因に対応するノードを障害ノードとして特定する。

【０１０８】

本実施の形態例における情報処理システムは、波及元となる割り込み要因に対応するノードを障害ノードとして特定することにより、複数の割り込み要因が連動して発生している場合に、複数の割り込み要因のうち、波及元の割り込み要因のみを対象として、当該波及元の割り込み要因に対応する障害ノードを特定することができる。

【０１０９】

また、本実施の形態例における情報処理システムにおいて、第１のノードの制御部は、障害として検出すべき割り込み要因を複数抽出した場合に、割り込み要因の優先度に基づいて、特定した障害ノードのメモリへのアクセスを抑止する。

【０１１０】

本実施の形態例における情報処理システムは、割り込み要因の優先度に基づいて、障害ノードのメモリへのアクセスの抑止処理の順を制御するため、割り込み要因に応じて、障害ノードのメモリへのアクセスの抑止処理の順を調整することができる。また、情報処理システムは、メモリへのアクセス抑止範囲の広い割り込み要因の優先度を低く設定することによって、情報処理システムの性能をより長く維持することができる。

【０１１１】
また、本実施の形態例における情報処理システムにおいて、１のノードの制御装置は、障害として検出すべき割り込み要因がデータ処理を実行するノードにおいて発生した割り込み要因である場合に、発生元のノードを障害ノードとして特定する。また、１のノードの制御装置は、障害として検出すべき割り込み要因が網結合装置を備えるノードにおいて発生した割り込み要因である場合に、網結合装置に接続されたノードを障害ノードとして特定する。このため、本実施の形態例における情報処理システムは、割り込み要因に基づいて、割り込み要因に対応する障害ノードを特定することができる。
【０１１２】
また、本実施の形態例における情報処理システムにおいて、１のノードは、割り込み要因と、割り込み要因の波及元となる割り込み要因との対応関係を有する定義テーブルを有し、１のノードの制御装置は、定義テーブルに基づいて、障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定する。
【０１１３】
本実施の形態例における情報処理システムは、割り込み要因と、割り込み要因の波及元となる割り込み要因との対応関係を有する定義テーブルを有することによって、波及元の割り込み要因であるか否かを高速に判定することができる。また、情報処理システムは、割り込み要因が増加した場合や、変更が発生した場合に、定義テーブルの更新処理を行うことで、割り込み要因の増加や変更を簡易に適用することができる。これにより、情報処理システムは、エンハンスや設計変更時におけるメンテナンス工数を小さく抑えることができる。
【０１１４】
また、本実施の形態例における情報処理システムは、割り込み要因に対応して優先度を有する定義テーブルを有し、１のノードの前記制御装置は、定義テーブルに基づいて、割り込み要因の優先度を判定する。
【０１１５】
本実施の形態例における情報処理システムは、割り込み要因に対応して優先度を有する定義テーブルを有することによって、割り込み要因の優先度を高速に取得することができる。また、情報処理システムは、割り込み要因が増加した場合や、変更が発生した場合に、定義テーブルの更新処理を行うことで、割り込み要因の増加や変更を簡易に適用することができる。これにより、情報処理システムは、エンハンスや設計変更時におけるメンテナンス工数を小さく抑えることができる。
【０１１６】
以上、各ノードが共有メモリを有する分散型共有メモリの構成を例に説明したが、本実施の形態は、各ノードが共有メモリを設けておらず、各ノードとは別に共有メモリを備えるクラスター型構成にも適用可能である。
【０１１７】
以上の実施の形態をまとめると、次の付記のとおりである。
【０１１８】
（付記１）
複数のノード間でメモリを共有する情報処理システムにおいて、
前記ノードの各々は、
複数の機能回路と前記機能回路を制御する制御装置と、
前記複数の機能回路から発生する割り込み要因を格納するレジスタとを有し、
前記複数のノードのうちの１のノードの前記制御装置は、
他の前記ノードの割り込み要因の発生に応じて前記レジスタの前記割り込み要因を受信し、前記割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、前記障害ノードの前記メモリへのアクセスを抑止後、前記他のノードから受信したログ情報に基づいて前記障害ノードの切り離し制御を行う情報処理システム。

【０１１９】
（付記２）
付記１において、
前記他のノードの前記制御装置は、前記レジスタの前記割り込み要因の発生を前記１のノードの制御装置に通知し、
前記１のノードの制御装置は、前記他のノードからの前記通知に応じて、前記他ノードの前記レジスタの割り込み要因と前記ログ情報とを収集する情報処理システム。
【０１２０】
（付記３）
付記１または２において、
前記１のノードは、網結合装置を備え、
前記他のノードは、データ処理を実行し、前記網結合装置を介して前記メモリにアクセスする処理装置を備える情報処理システム。
【０１２１】
（付記４）
付記１乃至３のいずれかにおいて、
前記１のノードの前記制御装置は、前記障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定し、発生していない場合に、前記割り込み要因に対応するノードを前記障害ノードとして特定し、発生している場合に、前記波及元となる割り込み要因に対応するノードを前記障害ノードとして特定する情報処理システム。
【０１２２】
（付記５）
付記１乃至４のいずれかにおいて、
前記第１のノードの前記制御部は、前記障害として検出すべき割り込み要因を複数抽出した場合に、前記割り込み要因の優先度に基づいて、前記特定した障害ノードの前記メモリへのアクセスを抑止する情報処理システム。
【０１２３】
（付記６）
付記３において、
前記１のノードの前記制御装置は、前記障害として検出すべき割り込み要因が前記データ処理を実行するノードにおいて発生した割り込み要因である場合に、発生元のノードを前記障害ノードとして特定し、前記障害として検出すべき割り込み要因が前記網結合装置を備えるノードにおいて発生した割り込み要因である場合に、前記網結合装置に接続されたノードを前記障害ノードとして特定する情報処理システム。
【０１２４】
（付記７）
付記４において、
前記１のノードは、
前記割り込み要因と、前記割り込み要因の波及元となる割り込み要因との対応関係を有する定義テーブルを有し、
前記１のノードの前記制御装置は、前記定義テーブルに基づいて、前記障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定する情報処理システム。
【０１２５】
（付記８）
付記５において、
前記１のノードは、
前記割り込み要因に対応して前記優先度を有する定義テーブルを有し、
前記１のノードの前記制御装置は、前記定義テーブルに基づいて、割り込み要因の前記

優先度を判定する情報処理システム。
【０１２６】
（付記９）
付記１乃至８のいずれかにおいて、
前記メモリは各前記ノード内に設けられた情報処理システム。
【０１２７】
（付記１０）
複数のノード間でメモリを共有する情報処理システムの障害処理方法において、
前記ノードの各々は、
複数の機能回路と前記機能回路を制御する制御装置と、
前記複数の機能回路から発生する割り込み要因を格納するレジスタとを有し、
前記複数のノードのうちの１のノードの前記制御装置は、
他の前記ノードの割り込み要因の発生に応じて前記レジスタの前記割り込み要因を受信し、前記割り込み要因のうち、障害として検出すべき割り込み要因を抽出して、抽出結果に応じて障害ノードを特定し、前記障害ノードの前記メモリへのアクセスを抑止後、前記他のノードから受信したログ情報に基づいて前記障害ノードの切り離し制御を行う情報処理システムの障害処理方法。
【０１２８】
（付記１１）
付記１０において、
前記他のノードの前記制御装置は、前記レジスタの前記割り込み要因の発生を前記１のノードの制御装置に通知し、
前記１のノードの制御装置は、前記他のノードからの前記通知に応じて、前記他ノードの前記レジスタの割り込み要因と前記ログ情報とを収集する情報処理システムの障害処理方法。
【０１２９】
（付記１２）
付記１０または１１において、
前記１のノードは、網結合装置を備え、
前記他のノードは、データ処理を実行し、前記網結合装置を介して前記メモリにアクセスする処理装置を備える情報処理システムの障害処理方法。
【０１３０】
（付記１３）
付記１０乃至１２のいずれかにおいて、
前記１のノードの前記制御装置は、前記障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定し、発生していない場合に、前記割り込み要因に対応するノードを前記障害ノードとして特定し、発生している場合に、前記波及元となる割り込み要因に対応するノードを前記障害ノードとして特定する情報処理システムの障害処理方法。
【０１３１】
（付記１４）
付記１０乃至１３のいずれかにおいて、
前記第１のノードの前記制御部は、前記障害として検出すべき割り込み要因を複数抽出した場合に、前記割り込み要因の優先度に基づいて、前記特定した障害ノードの前記メモリへのアクセスを抑止する情報処理システムの障害処理方法。
【０１３２】
（付記１５）
付記１２において、
前記１のノードの前記制御装置は、前記障害として検出すべき割り込み要因が前記データ処理を実行するノードにおいて発生した割り込み要因である場合に、発生元のノードを

前記障害ノードとして特定し、前記障害として検出すべき割り込み要因が前記網結合装置を備えるノードにおいて発生した割り込み要因である場合に、前記網結合装置に接続されたノードを前記障害ノードとして特定する情報処理システムの障害処理方法。

【０１３３】

（付記１６）

付記１３において、

前記１のノードは、

前記割り込み要因と、前記割り込み要因の波及元となる割り込み要因との対応関係を有する定義テーブルを有し、

前記１のノードの前記制御装置は、前記定義テーブルに基づいて、前記障害として検出すべき割り込み要因の波及元となる割り込み要因が発生しているか否かを判定する情報処理システムの障害処理方法。

【０１３４】

（付記１７）

付記１４において、

前記１のノードは、

前記割り込み要因に対応して前記優先度を有する定義テーブルを有し、

前記１のノードの前記制御装置は、前記定義テーブルに基づいて、割り込み要因の前記優先度を判定する情報処理システムの障害処理方法。

【０１３５】

（付記１８）

付記１０乃至１７のいずれかにおいて、

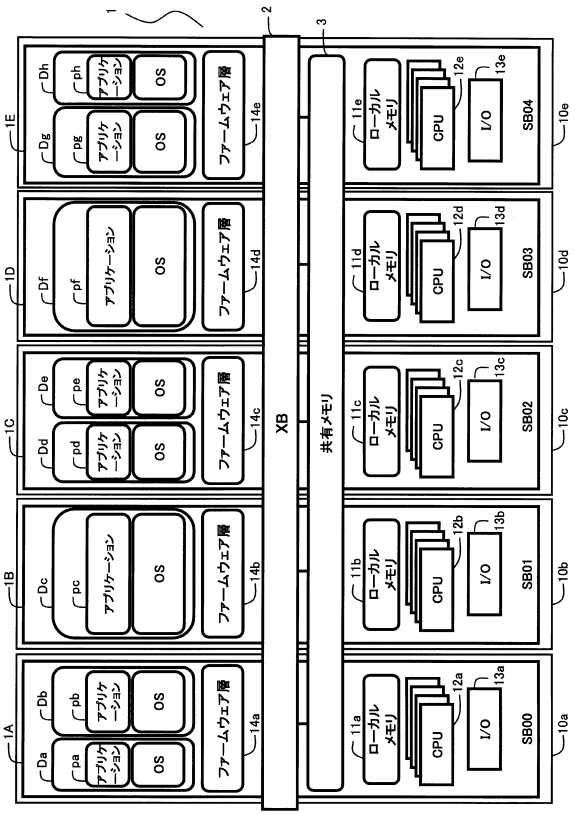
前記メモリは各前記ノード内に設けられた情報処理システムの障害処理方法。

【符号の説明】

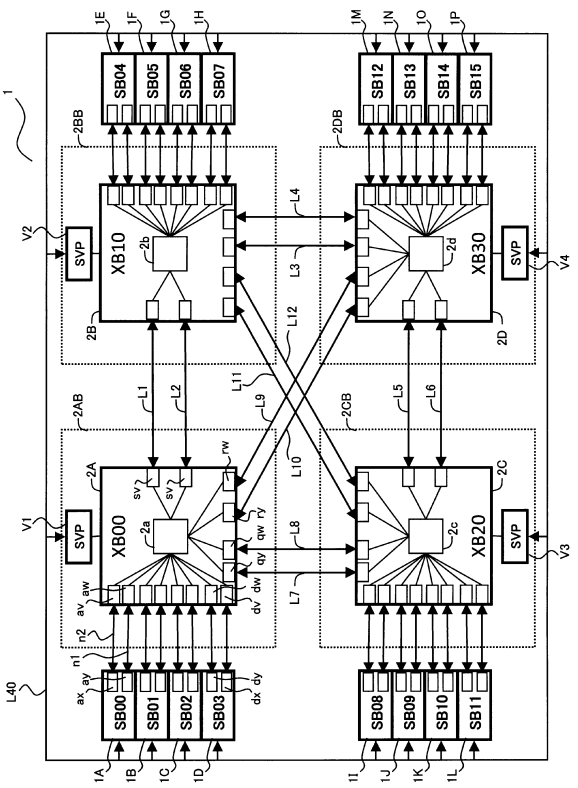
【０１３６】

１Ａ～１Ｐ：システムボード、２ＡＢ～２ＤＢ：クロスバスイッチボックス、
Ｂ１：システムボードユニット、１２：ＣＰＵチップ、１５：システムコントローラ、１６：Ｉ／Ｏコントローラ、１８：メモリコントローラ、１１：メモリ、１９：ＭＢＣ（システムボードユニット）、
Ｂ２：サービスプロセッサボード、ＭＢＣ２１（サービスプロセッサボードユニット）、
２２：システム制御装置、ｒｇ：レジスタ、
２ＡＢ：マスターノード、１Ｖ：システム制御装置、
１Ａ：スレーブノード、２２：システム制御装置

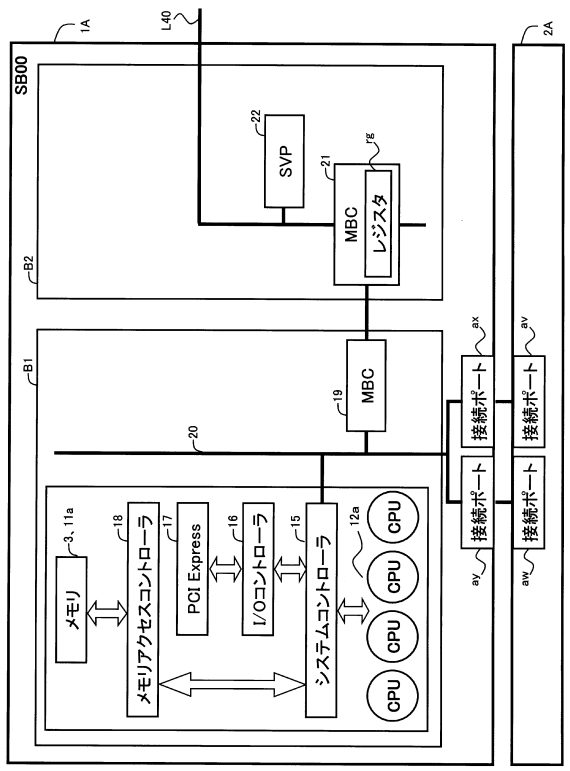
【図１】


1
2
3
XB
共有メモリ
SB00
SB01
SB02
SB03
SB04
CPU
I/O
ローカルメモリ
ファームウェア層
アプリケーション
OS

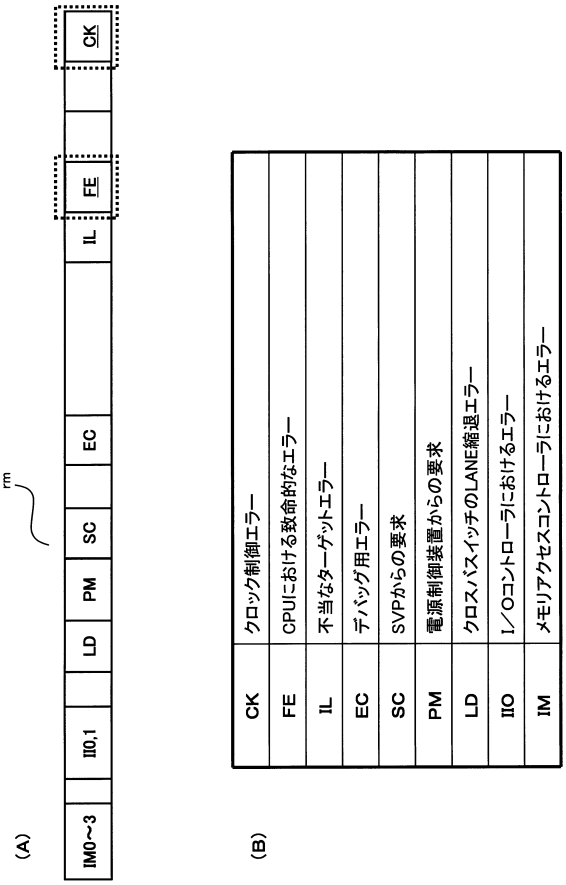
【図２】


XB00
XB10
XB20
XB30
SVP
SB00
SB01
SB02
SB03
SB04
SB05
SB06
SB07
SB08
SB09
SB10
SB11
SB12
SB13
SB14
SB15

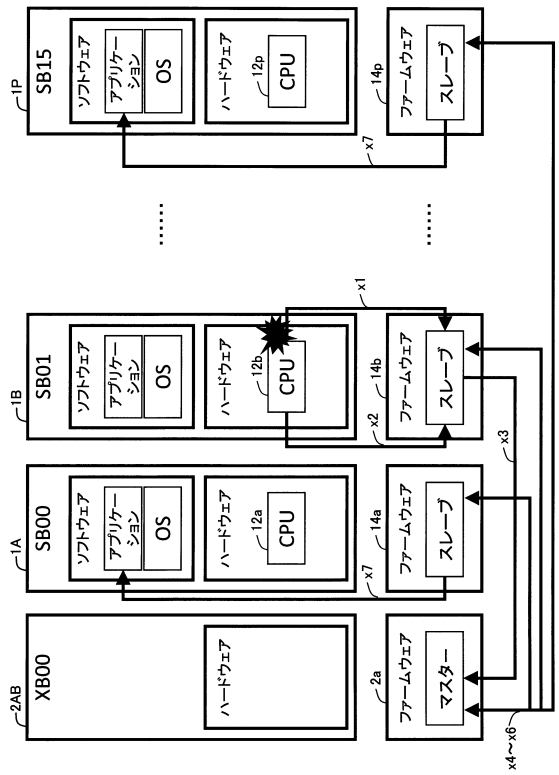
【図３】


SB00
SVP
MBC
レジスタ
メモリ
メモリアクセスコントローラ
PCI Express
I/Oコントローラ
システムコントローラ
CPU
接続ポート

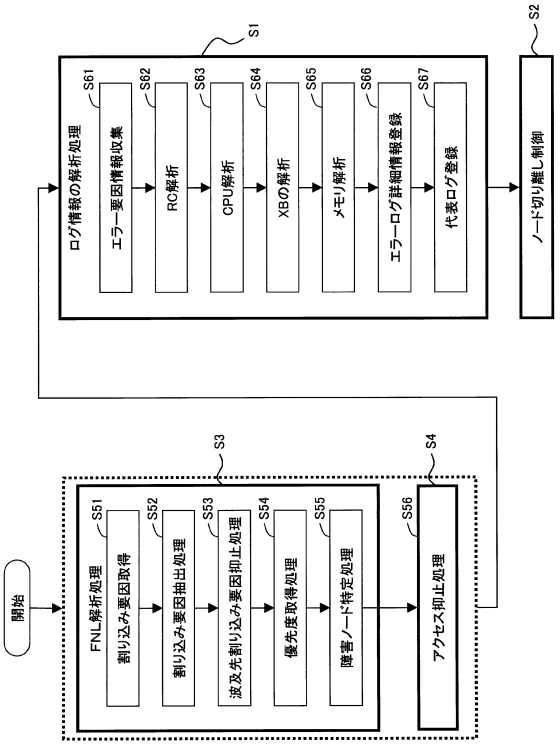
【図４】


(A)
IM0~3
IIO,1
LD
PM
SC
EC
IL
FE
CK
(B)
CK クロック制御エラー
FE CPUにおける致命的なエラー
IL 不当なターゲットエラー
EC デバッグ用エラー
SC SVPからの要求
PM 電源制御装置からの要求
LD クロスバスイッチのLANE縮退エラー
IIO I/Oコントローラにおけるエラー
IM メモリアクセスコントローラにおけるエラー

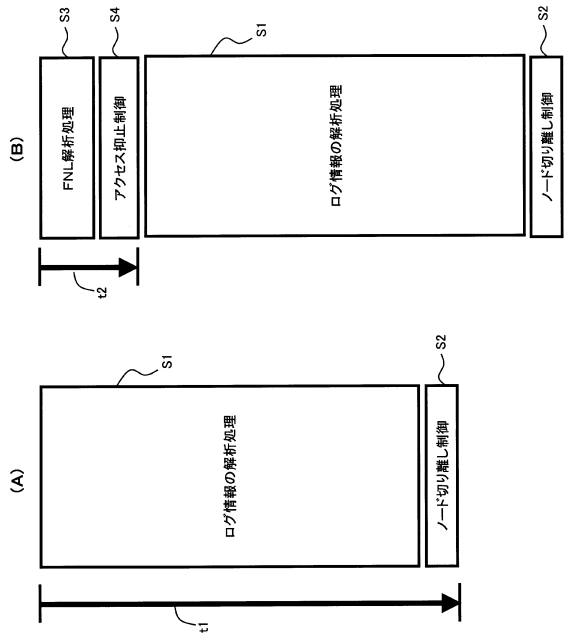
【図５】


1P
SB15
ソフトウェア
アプリケーション
OS
ハードウェア
12p
CPU
14p
ファームウェア
スレーブ
x7
1B
SB01
12b
14b
x1
x2
1A
SB00
12a
14a
x3
2AB
XB00
2a
マスター
x4～x6

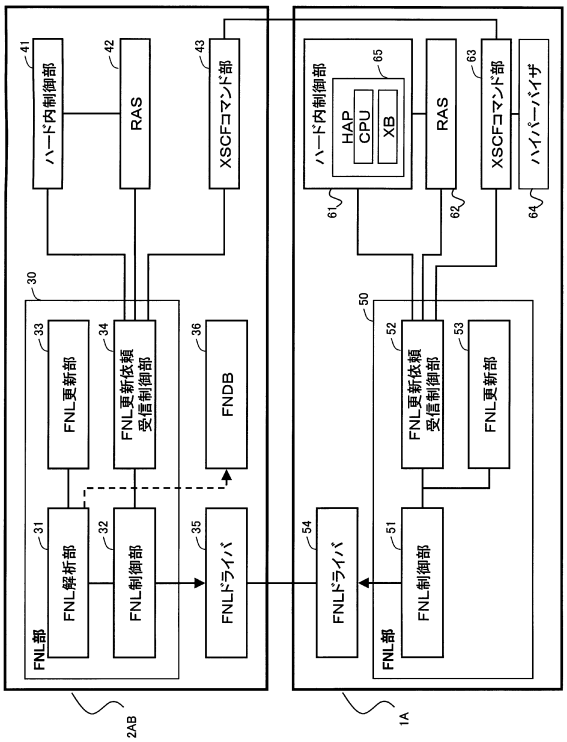
【図６】


S1
S2
ログ情報の解析処理
エラー要因情報収集
RC解析
CPU解析
XBの解析
メモリ解析
エラーログ詳細情報登録
代表ログ登録
ノード切り離し制御
開始
FNL解析処理
割り込み要因取得
割り込み要因抽出処理
波及先割り込み要因抑止処理
優先度取得処理
障害ノード特定処理
アクセス抑止処理
S3
S4
S51
S52
S53
S54
S55
S56

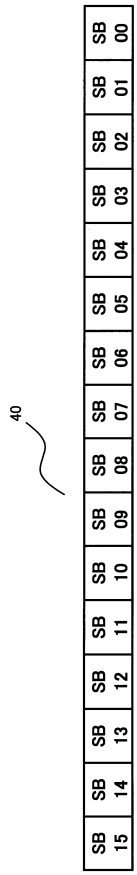
【図７】


(A)
(B)
S1
S2
S3
S4
ログ情報の解析処理
ノード切り離し制御
FNL解析処理
アクセス抑止制御
t1
t2

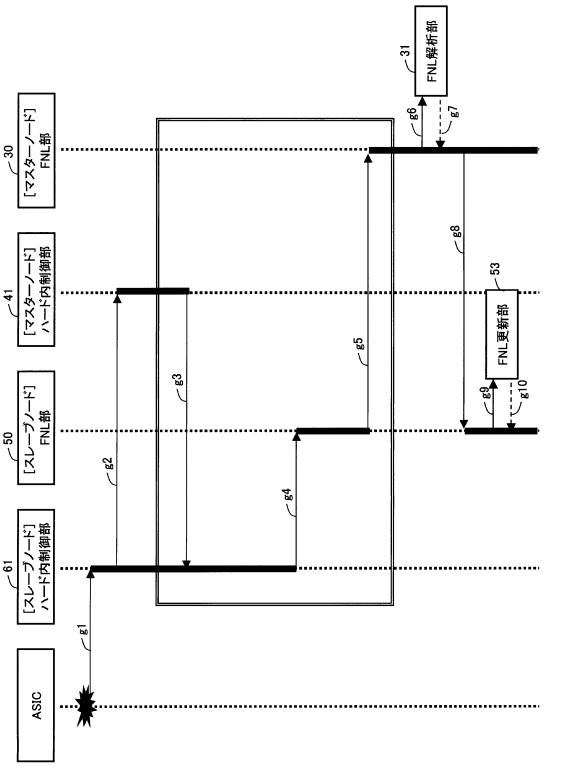
【図８】


FNL部
FNL解析部
FNL制御部
FNLドライバ
FNL更新部
FNL更新依頼受信制御部
FNDB
ハード内制御部
RAS
XSCFコマンド部
HAP
CPU
XB
ハイパーバイザ
2AB
1A
30
31
32
33
34
35
36
41
42
43
50
51
52
53
54
61
62
63
64
65

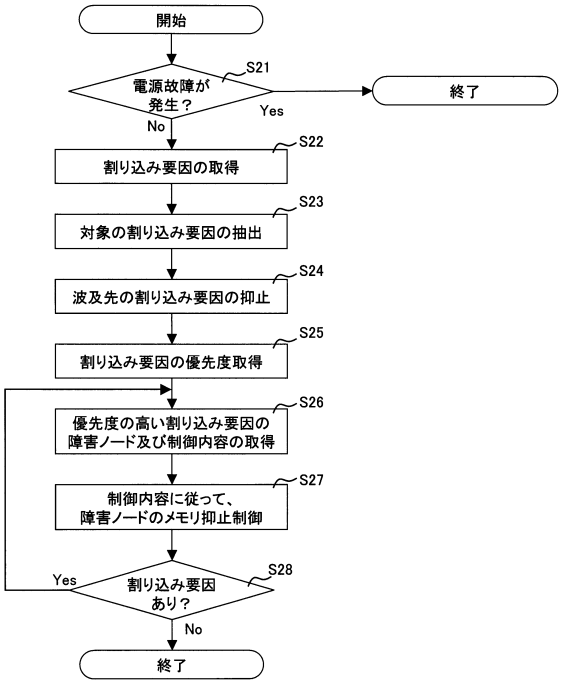
【図９】



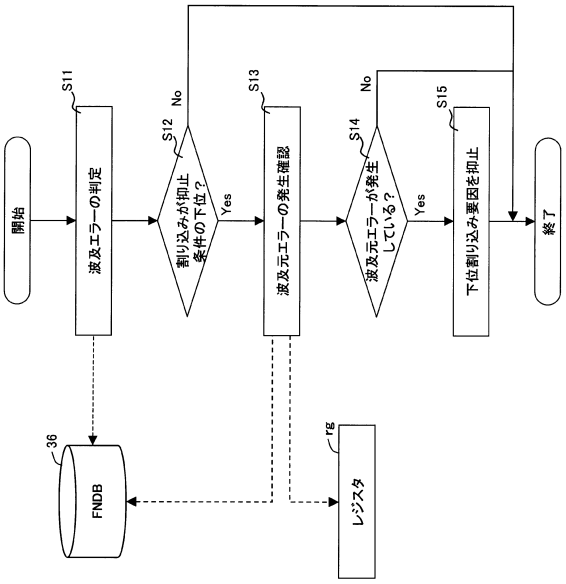
【図１０】



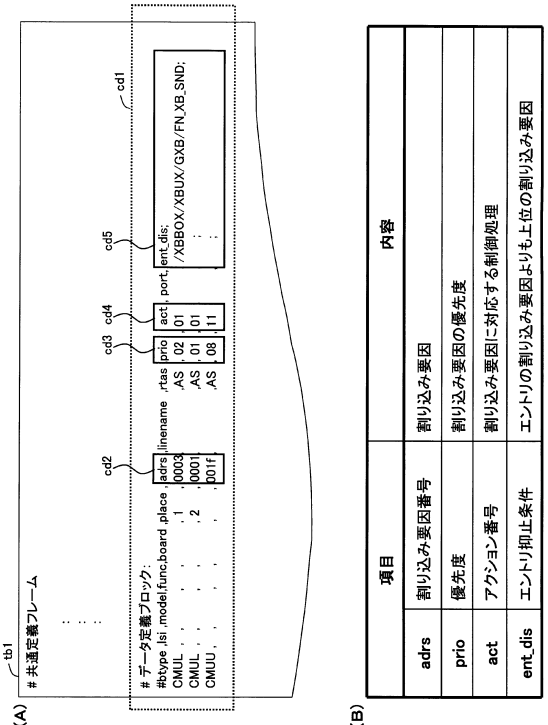
【図１１】



【図１２】



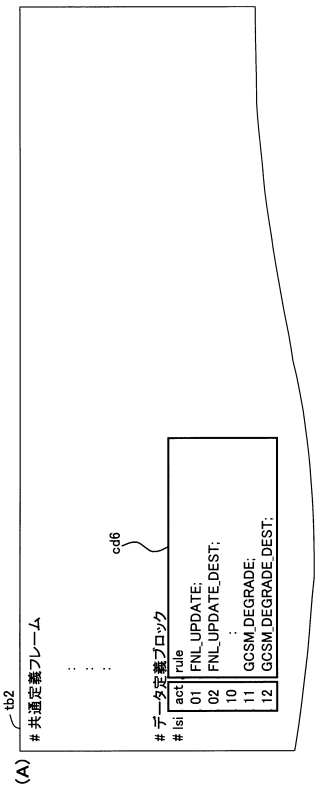
【図１３】



項目	内容
adrs	割り込み要因番号
prio	優先度
act	アクション番号
ent_dis	エントリ抑止条件

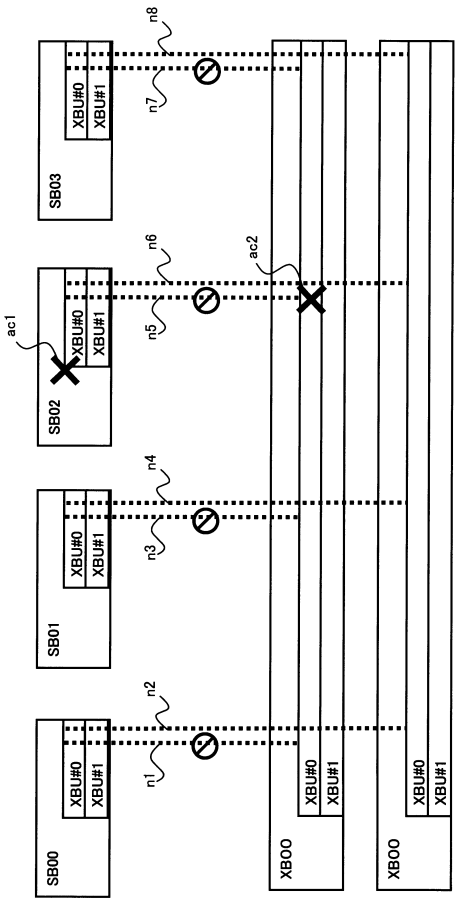
内容
割り込み要因
割り込み要因の優先度
割り込み要因に対応する制御処理
エントリの割り込み要因よりも上位の割り込み要因

【図１４】



定義(rule)		内容
ノードダウン	FNL_UPDATE	割り込み要因の検出ノードを停止制御
	FNL_UPDATE_DEST	割り込み要因の検出ノードに接続されたノードを停止制御
機能縮退	GCSM_DEGRADE	割り込み要因の検出ノードを機能縮退
	GCSM_DEGRADE_DEST	割り込み要因の検出ノードに接続されたノードを機能縮退

【図１５】



フロントページの続き

(56)参考文献 特開２００１－１３４５４６（ＪＰ，Ａ）
特開２０１３－１４０４４５（ＪＰ，Ａ）
特開２００６－１７８７８６（ＪＰ，Ａ）
特開２００４－０６２５３５（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１１／０７
Ｇ０６Ｆ　１１／３０－１１／３４