

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2014-6676

(P2014-6676A)

(43) 公開日 平成26年1月16日(2014.1.16)

(51) Int. Cl. F I テーマコード (参考)
G06F 11/34 (2006.01) G06F 11/34 Q 5B042
G06F 9/46 (2006.01) G06F 9/46 350

審査請求 未請求 請求項の数 17 O L (全 51 頁)

(21) 出願番号	特願2012-141495 (P2012-141495)	(71) 出願人	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
(22) 出願日	平成24年6月22日 (2012.6.22)	(74) 代理人	100074099 弁理士 大菅 義之
		(74) 代理人	100133570 弁理士 ▲徳▼永 民雄
		(72) 発明者	近藤 浩 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
		(72) 発明者	岡野 憲司 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
		Fターム(参考)	5B042 GA22 GA33 GC08 KK02 KK08 KK20 MB01 MC07

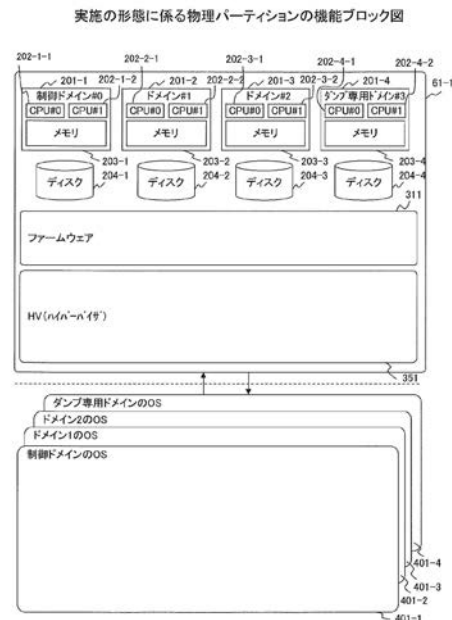
(54) 【発明の名称】 メモリダンプ機能を有する情報処理装置、メモリダンプ方法、およびメモリダンププログラム

(57) 【要約】

【課題】再起動が必要となる異常を検出した場合、速やかに業務を再開させて、原因調査のためのメモリダンプを採取すること。

【解決手段】ハイパーバイザのエラーまたはオペレーティングシステムが異常を検出したときに、前記オペレーティングシステムのカーネルまたは前記ハイパーバイザが使用している第1のメモリ領域をファームウェアに通知し、前記ハイパーバイザおよびオペレーティングシステムを停止し、前記オペレーティングシステムのカーネルまたは前記ハイパーバイザが使用するメモリ領域を前記第1のメモリ領域とは異なる第2のメモリ領域に変更し、前記ハイパーバイザおよび前記オペレーティングシステムを起動して業務を再開させ、前記第1のメモリ領域のデータを読み出して、該データをファイルに書き出す。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

メモリと、

前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステム、および前記仮想マシンを制御するハイパーバイザを実行する処理部と、

前記メモリ及び前記処理部を含むシステムの制御を行なう制御部と、

を有し、

前記処理部は、

前記ハイパーバイザのエラーを検出したときに、前記オペレーティングシステムを停止し、

前記ハイパーバイザが使用している第 1 のメモリ領域を前記制御部に通知し、

前記ハイパーバイザを停止し、

前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第 1 のメモリ領域とは異なる第 2 のメモリ領域に変更し、

前記第 2 のメモリ領域を使用領域として、前記ハイパーバイザを起動し、

前記オペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第 1 のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

ことを特徴とする情報処理装置。

【請求項 2】

前記処理部は、さらに

前記第 2 のメモリ領域を使用領域として前記ハイパーバイザを起動した後、停止した前記オペレーティングシステムが使用するメモリ領域を、前記オペレーティングシステムのカーネルが使用していた第 3 のメモリ領域とは異なる第 4 のメモリ領域に変更し、

前記第 4 のメモリ領域を使用領域として、停止した前記オペレーティングシステムを起動し、

前記第 3 のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出す

ことを特徴とする請求項 1 記載の情報処理装置。

【請求項 3】

前記処理部は、前記ハイパーバイザのエラーを検出したときに、前記第 1 のメモリ領域を、ダンプ対象領域として前記制御部に通知することを特徴とする、請求項 1 または 2 に記載の情報処理装置。

【請求項 4】

前記ハイパーバイザを停止した後、前記第 1 のメモリ領域の内容を保持したまま、前記ハイパーバイザが使用するメモリ領域を前記第 2 のメモリ領域に変更し、少なくとも変更後の前記第 2 のメモリ領域を初期化した後に、前記ハイパーバイザを起動する、ことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の情報処理装置。

【請求項 5】

さらに、前記ハイパーバイザが使用していた領域のダンプを行うか否かを示すフラグ情報を有し、

前記フラグ情報が前記ハイパーバイザが使用していた領域のダンプを行うことを示す場合、前記第 1 のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

ことを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の情報処理装置。

【請求項 6】

前記フラグ情報に基づいて、前記第 1 のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す他の仮想マシンをさらに実行

10

20

30

40

50

することを特徴とする、請求項 5 記載の情報処理装置。

【請求項 7】

前記データをハイパーバイザのダンプファイルとしてファイルに書き出した後に、前記フラグ情報を初期化することを特徴とする請求項 6 記載の情報処理装置。

【請求項 8】

前記データをハイパーバイザのダンプファイルとしてファイルに書き出した後に、前記第 1 のメモリ領域を資源として開放する、請求項 1 乃至 8 のいずれか 1 項に記載の情報処理装置。

【請求項 9】

メモリと、

オペレーティングシステムを実行、または、仮想マシン及び該仮想マシン上で稼動するオペレーティングシステムを実行する処理部と

を有し、

前記処理部は、

実行中のオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、

停止したオペレーティングシステムが使用するメモリ領域を、前記停止したオペレーティングシステムのカーネルが使用していた第 1 のメモリ領域とは異なる第 2 のメモリ領域に変更し、

前記第 2 のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第 1 のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出す

ことを特徴とする情報処理装置。

【請求項 10】

前記処理部は、

前記仮想マシン上で稼動するオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、

前記第 2 のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動し、

前記第 1 のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出すとともに、

前記仮想マシンを制御するハイパーバイザを稼動させたまま、前記ハイパーバイザが使用するメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

ことを特徴とする、請求項 9 記載の情報処理装置。

【請求項 11】

各々メモリと、前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステムを実行する処理部とを有する複数の物理パーティションと、

前記複数の物理パーティションの制御を行なう制御部と、

を有し、

前記物理ドメインの各々に含まれる前記処理部は、

実行中のオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、

停止したオペレーティングシステムが使用するメモリ領域を、前記停止したオペレーティングシステムのカーネルが使用していた第 1 のメモリ領域とは異なる第 2 のメモリ領域に変更し、

前記第 2 のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

10

20

30

40

50

前記第 1 のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出す

ことを特徴とする情報処理装置。

【請求項 1 2】

各々メモリと、前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステム、および前記仮想マシンを制御するハイパーバイザを実行する処理部とを有する複数の物理パーティションと、

前記複数の物理パーティションの制御を行なう制御部と、
を有し、

前記物理ドメインの各々に含まれる前記処理部は、

前記ハイパーバイザのエラーを検出したときに、前記オペレーティングシステムを停止し、

前記ハイパーバイザが使用している第 1 のメモリ領域を前記制御部に通知し、

前記ハイパーバイザを停止し、

前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第 1 のメモリ領域とは異なる第 2 のメモリ領域に変更し、

前記第 2 のメモリ領域を使用領域として、前記ハイパーバイザを起動し、

前記オペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第 1 のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

ことを特徴とする情報処理装置。

【請求項 1 3】

前記物理パーティションの各々では、複数の仮想マシンを実行し、

前記物理パーティションの各々に含まれる前記処理部は、

前記ハイパーバイザのエラーを検出したときに、複数の仮想マシンの各々で稼動するオペレーティングシステムを停止し、

前記第 1 のメモリ領域を前記制御部に通知し、

前記ハイパーバイザを停止し、

前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第 1 のメモリ領域とは異なる第 2 のメモリ領域に変更し、

前記第 2 のメモリ領域を使用領域として前記ハイパーバイザを起動した後、停止した複数のオペレーティングシステムの各々が使用するメモリ領域を、前記停止した複数のオペレーティングシステムの各々のカーネルが使用していた複数の第 3 のメモリ領域とは異なり、各々重複しない複数の第 4 のメモリ領域に変更し、

前記複数の第 4 のメモリ領域の各々を使用領域として、各使用領域に対応する前記停止したオペレーティングシステムの各々を起動して、起動した各オペレーティングシステム上で稼動する業務プログラムを再開し、

前記複数の第 3 のメモリ領域の各々データを読み出して、該データを前記複数のオペレーティングシステムのダンプファイルとしてファイルに書き出す

ことを特徴とする請求項 1 2 記載の情報処理装置。

【請求項 1 4】

メモリと、前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステム、および前記仮想マシンを制御するハイパーバイザを実行する処理部と、前記メモリおよび前記処理部を含むシステムの制御を行う制御部と、を有する情報処理装置が実行するメモリダンプ方法であって、

前記ハイパーバイザのエラーを検出したときに、前記オペレーティングシステムを停止し、

前記ハイパーバイザが使用している第 1 のメモリ領域を前記制御部に通知し、

前記ハイパーバイザを停止し、

10

20

30

40

50

前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第1のメモリ領域とは異なる第2のメモリ領域に変更し、

前記第2のメモリ領域を使用領域として、前記ハイパーバイザを起動し、

前記オペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第1のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

ことを特徴とするメモリダンプ方法。

【請求項15】

オペレーティングシステムを実行、または、仮想マシン及び該仮想マシン上で稼動するオペレーティングシステムを実行する情報処理装置が実行するメモリダンプ方法であって、

10

実行中のオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、

停止したオペレーティングシステムが使用するメモリ領域を、前記停止したオペレーティングシステムのカーネルが使用していた第1のメモリ領域とは異なる第2のメモリ領域に変更し、

前記第2のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第1のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出す

20

ことを特徴とするメモリダンプ方法。

【請求項16】

メモリと、前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステム、および前記仮想マシンを制御するハイパーバイザを実行する処理部と、前記メモリおよび前記処理部を含むシステムの制御を行う制御部と、を有するコンピュータに、

前記ハイパーバイザのエラーを検出したときに、前記オペレーティングシステムを停止し、

前記ハイパーバイザが使用している第1のメモリ領域を前記制御部に通知し、

30

前記ハイパーバイザを停止し、

前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第1のメモリ領域とは異なる第2のメモリ領域に変更し、

前記第2のメモリ領域を使用領域として、前記ハイパーバイザを起動し、

前記オペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第1のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す

処理を実行させるメモリダンププログラム。

【請求項17】

40

オペレーティングシステムを実行、または、仮想マシン及び該仮想マシン上で稼動するオペレーティングシステムを実行するコンピュータに、

実行中のオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、

停止したオペレーティングシステムが使用するメモリ領域を、前記停止したオペレーティングシステムのカーネルが使用していた第1のメモリ領域とは異なる第2のメモリ領域に変更し、

前記第2のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、

前記第1のメモリ領域のデータを読み出して、該データをオペレーティングシステムの

50

ダンプファイルとしてファイルに書き出す

処理を実行させるメモリダンププログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、メモリダンプ機能を有する情報処理装置、メモリダンプ方法、およびメモリダンププログラムに関する。

【背景技術】

【0002】

近年、UNIX（登録商標）サーバ、IAサーバが基幹システムに導入されるようになり、UNIX（登録商標）サーバ、IAサーバの高可用性が重要視されている。一般的に、システムに致命的なエラーが発生した場合はシステムを緊急停止（パニック）させて、その原因を調査するためにメモリダンプをディスクに保存している。 10

【0003】

システムを停止している間は、システムを使用できないので、速やかにシステムを再起動することが重要な要件となる。

しかし、近年では、実装メモリの容量がテラバイト（TB）オーダのサーバが登場し、このようなシステムでは、メモリダンプを採取するのに非常に時間がかかり、速やかにシステムを再起動することができなくなっている。

【0004】

また、メモリダンプをディスク上に保存せず、緊急停止時のメモリ内容を別のメモリ上に保存する方法や障害発生時のメモリ内容をダンプ格納領域に保存する際、メモリの一部を保存し、再起動後に保存していないメモリ内容をダンプファイルに変換する方法が知られている。 20

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開平11-212836号公報

【特許文献2】特開2001-229053号公報

【特許文献3】特開2006-72931号公報 30

【特許文献4】特開2005-122334号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、従来の方法では、異常発生時のメモリダンプを別のメモリやディスクに保存しているため、保存するメモリダンプのサイズが大きい場合は、メモリのコピーに時間がかかり、速やかにシステムを再起動することができないという問題があった。

【0007】

また、オペレーティングシステムが致命的なエラーを検出して、システムを緊急停止する場合、異常を検出したオペレーティングシステムがダンプを採取するため、ダンプ採取処理中に再度異常を検出して、ハングアップが発生するなどの二次被害が発生することがあるという問題があった。 40

【0008】

1つの側面では、本発明は、再起動が必要となる異常を検出した場合、速やかに業務を再開させて、原因調査のためのメモリダンプを採取することを課題とする。

【課題を解決するための手段】

【0009】

1つの側面では、実施の形態の情報処理装置は、メモリと、前記メモリに格納されたプログラムを実行することにより、仮想マシン、該仮想マシン上で稼動するオペレーティングシステム、および前記仮想マシンを制御するハイパーバイザを実行する処理部と、前記 50

メモリ及び前記処理部を含むシステムの制御を行なう制御部と、を有する。

【0010】

前記処理部は、前記ハイパーバイザのエラーを検出したときに、前記オペレーティングシステムを停止し、前記ハイパーバイザが使用している第1のメモリ領域を前記制御部に通知する。

【0011】

そして、前記処理部は、前記ハイパーバイザを停止し、前記ハイパーバイザが使用するメモリ領域を、前記制御部に通知された前記第1のメモリ領域とは異なる第2のメモリ領域に変更し、前記第2のメモリ領域を使用領域として、前記ハイパーバイザを起動する。

【0012】

さらに、前記処理部は、前記オペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、前記第1のメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出す。

【0013】

他の1つの側面では、実施の形態の情報処理装置は、メモリと、オペレーティングシステムを実行、または、仮想マシン及び該仮想マシン上で稼動するオペレーティングシステムを実行する処理部と、を有する。

【0014】

前記処理部は、実行中のオペレーティングシステムのエラーを検出したときに、前記エラーを検出したオペレーティングシステムを停止し、停止したオペレーティングシステムが使用するメモリ領域を、前記停止したオペレーティングシステムのカーネルが使用していた第1のメモリ領域とは異なる第2のメモリ領域に変更する。

【0015】

そして、前記処理部は、前記第2のメモリ領域を使用領域として、前記停止したオペレーティングシステムを起動して該オペレーティングシステム上で稼動する業務プログラムを再開し、前記第1のメモリ領域のデータを読み出して、該データをオペレーティングシステムのダンプファイルとしてファイルに書き出す。

【発明の効果】

【0016】

1つの実施の形態の情報処理装置によれば、再起動が必要となる異常を検出した場合、速やかに業務を再開させて、原因調査のためのメモリダンプを採取することができる。

【図面の簡単な説明】

【0017】

【図1】実施の形態に係るサーバのハードウェア構成図である。

【図2】実施の形態に係るサーバと機能との対応関係を示す図である。

【図3】実施の形態に係る物理パーティションの機能ブロック図である。

【図4】実施の形態に係るファームウェアの構成図である。

【図5】実施の形態に係るハイパーバイザの構成図である。

【図6】実施の形態に係るOSの構成図である。

【図7A】第1の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図7B】第1の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図7C】第1の実施の形態に係るメモリダンプ生成処理の変形例のフローチャートである。

【図8】HVダンプ対象領域情報の例である。

【図9A】第2の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図9B】第2の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図9C】第2の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図10】カーネルダンプ対象領域情報の例である。

【図11】ダンプ採取用ドメインによるカーネルのメモリダンプ生成処理のフローチャートである。

10

20

30

40

50

【図 1 2】 ダンプ採取用ドメインによるメモリダンプの採取を示す図である。

【図 1 3】 ダンプ採取用ドメインによるメモリダンプの採取におけるPA-RAマッピング情報を示す図である。

【図 1 4】 メモリDynamic Reconfiguration機能を用いたカーネルのメモリダンプ生成処理のフローチャートである。

【図 1 5】 メモリDynamic Reconfiguration機能を用いたメモリダンプの採取を示す図である。

【図 1 6】 メモリDynamic Reconfiguration機能を用いたメモリダンプの採取におけるPA-RAマッピング情報を示す図である。

【図 1 7】 第 3 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図 1 8 A】 第 4 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図 1 8 B】 第 4 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

【図 1 9】 稼動中のハイパーバイザのメモリダンプ生成処理のフローチャートである。

【発明を実施するための形態】

【0018】

以下、図面を参照しながら本発明の実施の形態を説明する。

図 1 は、実施の形態に係るサーバのハードウェア構成図である。

サーバ（情報処理装置）10 は、システムボード 11 - i（ $i = 1 \sim 3$ ）、サービスプロセッサ（SP）21、ディスクユニット 31、および通信インタフェース 41 を備える。

【0019】

システムボード 11 - i、サービスプロセッサ 21、ディスクユニット 31、および通信インタフェース 41 は、バス 51 を介して接続されている。

システムボード 11 - i は、Central Processing Unit（CPU）12 - i - k（ $k = 1, 2$ ）、不揮発性メモリ 14 - i、およびメモリ 13 - i - k を備える。

【0020】

サービスプロセッサ 21 は、サーバ 10 の制御、サーバ 10 内の物理パーティションの制御等を行う装置である。サービスプロセッサ 21 は、CPU 22 およびメモリ 23 を備える。サービスプロセッサ 21 は、制御部の一例である。

【0021】

CPU 22 は、サーバ 10 の制御、サーバ 10 内の物理パーティションの制御等の各種処理を行う。

メモリ 23 は、サービスプロセッサ 21 で用いられるデータを一時的に格納する。メモリ 23 は、例えば、RAM である。

【0022】

ディスクユニット 31 は、ハードディスクドライブ（HDD）32 - i（ $i = 1 \sim 3$ ）を備える。

HDD 32 は、サーバ 10 で使用されるデータを格納する装置である。HDD 32 は、記憶手段の一例である。

【0023】

通信インタフェース 41 は、サーバ 10 と接続する装置と通信を行うインタフェースである。

【0024】

図 2 は、実施の形態に係るサーバと機能との対応関係の一例を示す図である。

サーバ 10 は、2 つの物理パーティション 61 - k（ $k = 1, 2$ ）に分割して運用されている。尚、明細書内において、物理パーティション 61 - 1、61 - 2 をそれぞれ物理パーティション # 0、物理パーティション # 1 と表記する場合がある。

【0025】

物理パーティション # 0、# 1 は、サービスプロセッサ 21 により制御される。物理パーティション # 0、# 1 に含まれる CPU は、処理部の一例である。

10

20

30

40

50

システムボード 11-1、11-2 から構成される物理パーティション # 0 は、さらに物理パーティション # 0 内で 4 つの論理ドメイン # 0 ~ # 3 に分割して、各論理ドメイン # 0 ~ # 3 で独立したオペレーティングシステム (OS) が稼動している。また、ハイパーバイザ (HV) # 0 が、物理パーティション # 0 内の物理リソースと各論理ドメイン # 0 ~ # 3 との対応関係を制御する。

【0026】

システムボード 11-3 から構成される物理パーティション # 1 内では、論理ドメイン # 4 でオペレーティングシステム (OS) が稼動している。またハイパーバイザ # 1 が、物理パーティション # 1 内の物理リソースと論理ドメイン # 4 との対応関係を制御する。

【0027】

図 3 は、実施の形態に係る物理パーティションの機能ブロック図である。

物理パーティション 61-1 は、論理ドメイン 201-m (m = 1 ~ 4)、ファームウェア (FW) 311、およびハイパーバイザ (HV) 351 を備える。

【0028】

図 3 の物理パーティション 61-1 は、図 2 の物理パーティション 61-1 に対応する。

尚、論理ドメイン 201-1 ~ 201-4 は、それぞれ図 2 で示した各論理ドメイン # 0 ~ # 3 に対応する。

【0029】

尚、明細書内において、論理ドメイン 201-1 は、制御ドメイン # 0 と表記する場合がある。

また、明細書内において、論理ドメイン 201-4 は、ダンプ専用ドメイン # 3 またはダンプ採取用ドメイン 201-4 と表記する場合がある。

【0030】

以下、特に限定ない限り論理ドメインとは仮想マシンを示す。

論理ドメイン 201-m は、CPU 202-m-k (k = 1, 2)、メモリ 203-m、およびディスク 204-m を備える。以下、特に限定ない限り CPU 202、メモリ 203、およびディスク 204 は、それぞれ仮想 CPU、仮想メモリ、および仮想ディスクである。

【0031】

CPU 202-m-k は、各種処理を実行する。

メモリ 203-m は、ディスク 204-m から読み出された、各種プログラムやデータを格納する。

【0032】

ファームウェア 311 は、サーバ 10 全体 (複数の物理パーティション 61-1、61-2) の制御を行い、例えば、ハードウェアの初期化、メモリ診断、温度監視などを行う。ファームウェア 311 には、図 1 のサービスプロセッサ 21、およびシステムボード 11-1、11-2 の不揮発性メモリ 14-1、14-2 上に展開される Power On Self Test (POST) が含まれる。ここで、Power On Self Test (POST) とは、システム起動時にハードウェアリソースの診断と初期化を実行するプログラムである。

【0033】

ハイパーバイザ 351 は、論理ドメイン 201-m、および論理ドメイン 201-m 上で稼動するオペレーティングシステム (OS) 401-m を制御する。ハイパーバイザ 351 は、図 1 のシステムボード 11-1、11-2 のメモリ 13-1-1、13-1-2、13-2-1、13-2-2 上に展開され CPU 12-1-1、12-1-2、12-2-1、12-2-2 により実行される。

【0034】

図 3 の下部は、物理パーティション 61-1 内のソフトウェアを示す。

物理パーティション 61-1 内の各論理ドメイン # 0 ~ # 3 で、OS 401-m が稼動している。

10

20

30

40

50

OS 4 0 1 - 1 ~ 4 0 1 - 4 は、それぞれ論理ドメイン 2 0 1 - 1 ~ 2 0 1 - 4 のオペレーティングシステムに対応する。

【 0 0 3 5 】

図 4 は、実施の形態に係るファームウェアの詳細な構成図である。

ファームウェア 3 1 1 は、ダンプ対象領域情報/HVダンプフラグ格納処理部 3 1 2、ダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3、HVダンプフラグ設定部 3 1 4、メモリ初期化処理部 3 1 5、HV使用領域変更部 3 1 6、HV再起動命令部 3 1 7、ダンプ対象領域情報/HVダンプフラグ通知部 3 1 8、PA-RAマッピング通知部 3 1 9、メモリ開放処理部 3 2 0、HVダンプフラグリセット処理部 3 2 1 を備える。

【 0 0 3 6 】

ダンプ対象領域情報/HVダンプフラグ格納処理部 3 1 2 は、ダンプ対象領域情報およびHVダンプフラグをダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3 に格納する。

ダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3 は、ダンプ対象領域情報およびHVダンプフラグが格納される領域である。ここでダンプ対象領域情報は、ダンプ対象領域を示す情報であり、ダンプ対象領域の開始アドレス (PA Base) およびサイズの情報を含む。HVダンプフラグは、ハイパーバイザが使用していたメモリ領域のダンプファイルを生成するか否かを示す制御情報である。また、HVダンプフラグは、稼動中のハイパーバイザのメモリダンプを採取するか否かを示す情報 (HVライブダンプフラグ) を含むこともできる。

【 0 0 3 7 】

HVダンプフラグ設定部 3 1 4 は、HVダンプフラグの値を設定する。例えば、HVダンプフラグ設定部 3 1 4 は、HVダンプを行なう場合にHVダンプフラグをTRUEに設定する。

メモリ初期化処理部 3 1 5 は、メモリの初期化を行う。

【 0 0 3 8 】

HV使用領域変更部 3 1 6 は、ハイパーバイザ 3 5 1 が使用するメモリの領域を変更する。

HV再起動命令部 3 1 7 は、ハイパーバイザ 3 5 1 に再起動を指示する。

【 0 0 3 9 】

ダンプ対象領域情報/HVダンプフラグ通知部 3 1 8 は、ダンプ対象領域情報およびHVダンプフラグを通知する。

PA-RAマッピング通知部 3 1 9 は、OSの処理に必要なPA-RAマッピングをハイパーバイザ 3 5 1 へ通知し、ハイパーバイザ 3 5 1 のPA-RAマッピング処理部 3 6 8 は、通知されたPA-RAマッピングを用いて物理アドレス(PA)から実アドレス(RA)への、あるいは実アドレス(RA)から物理アドレス(PA)への変換を行う。

【 0 0 4 0 】

メモリ開放処理部 3 2 0 は、メモリの開放処理を行う。

HVダンプフラグリセット処理部 3 2 1 は、HVダンプフラグをリセットする。詳細には、HVダンプフラグリセット処理部 3 2 1 は、HVダンプフラグをFALSEに設定する。

【 0 0 4 1 】

図 5 は、実施の形態に係るハイパーバイザの詳細な構成図である。

ハイパーバイザ 3 5 1 は、ドメイン緊急停止指示部 3 5 2、OSパニック指示部 3 5 3、HVダンプ対象領域通知処理部 3 5 4、HVダンプ対象領域情報/HVダンプフラグ格納処理部 3 5 5、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6、HV再起動処理部 3 5 7、OS再起動命令部 3 5 8、HVメモリダンプフラグ読出・送信部 3 5 9、HVダンプ対象領域読出処理部 3 6 0、メモリ管理部 3 6 1、メモリ開放処理部 3 6 2、HVダンプフラグリセット処理部 3 6 3、HVダンプフラグ通知部 3 6 4、ダンプ専用ドメイン起動処理部 3 6 5、カーネルダンプ対象領域情報/カーネルダンプフラグ格納処理部 3 6 6、カーネルダンプ対象領域情報/カーネルダンプフラグ格納領域 3 6 7、PA-RAマッピング処理部 3 6 8、PA-RAマッピング情報格納域 3 6 9、割り込み処理部 3 7 0、メモリダンプ処理起動部 3 7 1、メモリ初期化処理部 3 7 2、およびカーネルダンプフラグリセット処理部 3 7 3 を

10

20

30

40

50

備える。

【 0 0 4 2 】

ドメイン緊急停止指示部 3 5 2 は、ドメイン 2 0 1 に緊急停止を指示する。

OSパニック指示部 3 5 3 は、OS 4 0 1 に緊急停止（パニック）指示する。

HVダンプ対象領域通知処理部 3 5 4 は、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 からHVダンプ対象領域情報を読み出して通知する。

【 0 0 4 3 】

HVダンプ対象領域情報/HVダンプフラグ格納処理部 3 5 5 は、HVダンプ対象領域情報およびHVダンプフラグをHVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 に格納する。

10

【 0 0 4 4 】

HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 は、HVダンプ対象領域情報およびHVダンプフラグを格納する。HVダンプ対象領域情報は、ハイパーバイザ 3 5 1 が使用しているメモリ領域（HVダンプ対象領域）を示す情報であり、メモリ領域の開始アドレス（PA Base）およびサイズの情報を含む。HVダンプフラグは、ハイパーバイザが使用していたメモリ領域のダンプファイルを生成するか否かを示す制御情報である。

【 0 0 4 5 】

HV再起動処理部 3 5 7 は、ハイパーバイザ 3 5 1 を停止させ、ハイパーバイザ 3 5 1 の再起動を行う。

OS再起動司令部 3 5 8 は、OS 4 0 1 に再起動を指示する。

20

【 0 0 4 6 】

HVメモリダンプフラグ読出・送信部 3 5 9 は、HVダンプフラグを読み出して送信する。

HVダンプ対象領域読出処理部 3 6 0 は、HVダンプ対象領域情報で示されるメモリ領域の内容を読み出し、送信する。または、HVダンプ対象領域読出処理部 3 6 0 は、現在のハイパーバイザ 3 5 1 が使用しているメモリ領域の内容を読み出し、送信する。

【 0 0 4 7 】

メモリ管理部 3 6 1 は、メモリを管理する。

メモリ開放処理部 3 6 2 は、メモリの開放処理を行う。

HVダンプフラグリセット処理部 3 6 3 は、HVダンプフラグをリセットする。詳細には、HVダンプフラグリセット処理部 3 6 3 は、例えば、HVダンプフラグをFALSEに設定する。

30

【 0 0 4 8 】

HVダンプフラグ通知部 3 6 4 は、HVダンプフラグを通知する。

ダンプ専用ドメイン起動処理部 3 6 5 は、ダンプ専用ドメインをファームウェアモードで起動する。ファームウェアモードとは、OSを起動しないモード、すなわちOSを起動する前に停止するモードである。

【 0 0 4 9 】

カーネルダンプ対象領域情報/カーネルダンプフラグ格納処理部 3 6 6 は、カーネルダンプ対象領域情報およびカーネルダンプフラグをカーネルダンプ対象領域情報/カーネルダンプフラグ格納領域 3 6 7 に格納する。

【 0 0 5 0 】

40

カーネルダンプ対象領域情報/カーネルダンプフラグ格納領域 3 6 7 は、カーネルダンプ対象領域情報およびカーネルダンプフラグを格納する。カーネルダンプ対象領域情報は、パニック時にOS 4 0 1 のカーネルが使用していたメモリ領域（カーネルダンプ対象領域）を示す情報であり、メモリ領域の開始アドレス（RA Base）およびサイズの情報を含む。カーネルダンプフラグは、OS 4 0 1 のカーネルのメモリダンプを実行するか否かを示す情報である。さらに、カーネルダンプフラグは、どのような方法でカーネルのメモリダンプを採取するかを示すこともできる。カーネルダンプフラグは、例えば、0：カーネルのメモリダンプを採取しない、1：ダンプ採取用ドメインで採取、または2：メモリDynamic Reconfiguration機能を用いて採取というような情報である。カーネルダンプフラグは、OS 4 0 1 から受信しても良いし、ハイパーバイザ 3 5 1 が予め設定して保持していても

50

良い。

【0051】

PA-RAマッピング処理部368は、物理アドレス(PA)とリアルアドレス(RA)間のマッピングを行う。PAはメモリの物理アドレスであり、RAはドメイン(オペレーティングシステム)上の実アドレスである。

【0052】

PA-RAマッピング情報格納域369は、PAとRA間のマッピングの情報を格納する。

割り込み処理部370は、OS401のカーネルのメモリダンプをする場合にS401に対して、割り込み処理させる。割り込み処理が受け付けられるとOS401のカーネルのメモリダンプが可能と判断されてOS401のカーネルのメモリダンプ処理に進み、受け付けられないと割り込み出来ないと判断してOS401のカーネルのメモリダンプ処理はせずに終了する。

10

【0053】

メモリダンプ処理起動部371は、制御ドメイン201-1にハイパーバイザ351のメモリダンプ処理を起動させる。

メモリ初期化処理部372は、メモリの初期化を行う。

【0054】

カーネルダンプフラグリセット処理部373は、カーネルダンプフラグをリセットする。例えば、カーネルダンプフラグリセット処理部373は、カーネルダンプフラグを削除または“0:カーネルのメモリダンプを採取しない”に設定する。

20

【0055】

図6は、実施の形態に係るOSの詳細な構成図である。

OS401-mは、メモリ管理部402-m、ファイル管理部403-m、プロセス管理部404-m、割り込み処理部405-m、マッピング情報抽出・格納処理部406-m、マッピング情報格納領域407-m、HVメモリダンプ判断部408-m、OS起動処理部409-m、HVダンプ対象領域読出処理部の呼び出し部410-m、カーネルダンプ対象メモリ読出処理部411-m、HVダンプ採取処理部412-m、カーネルダンプ採取処理部413-m、パニック処理部414-m、カーネルダンプ対象領域通知処理部415-m、メモリDR組み込み処理部416-m、メモリDR切り離し処理部417-m、空きメモリ追加処理部418-m、ダンプ専用ドメイン停止処理部419-m、カーネルダンプフラグリセット処理部420-m、およびカーネルダンプフラグ通知部421-mを備える。

30

【0056】

メモリ管理部402-mは、OS401-mが使用するメモリ203-mを割り当てる。

ファイル管理部403-mは、ディスク上に格納されたデータであるファイルを管理する。

【0057】

プロセス管理部404-mは、OS401-mが実行するプログラムのプロセスを管理する。

割り込み処理部405-mは、割り込み処理を行う。

40

【0058】

マッピング情報抽出・格納処理部406-mは、メモリ203-mのダンプを取得および解析するために必要な情報を、マッピング情報格納領域407-mに格納する。

マッピング情報格納領域407-mは、メモリ203-mのダンプを取得および解析するために必要な情報を格納する。マッピング情報格納領域407-mが格納する情報は、例えば、カーネルのテキスト域、データ域、ヒープ域、スタック域等、各セグメントのマッピング情報(論理アドレス、物理アドレス、サイズ等)やアドレス変換テーブル、ページテーブル等、各種制御テーブルのマッピング情報である。

【0059】

HVメモリダンプ判断部408-mは、HVダンプフラグがTRUEまたはFALSEであるか判定

50

し、ハイパーバイザのメモリダンプを行うか判断する。

OS起動処理部 4 0 9 - m は、OS 4 0 1 - m を再起動する。

【 0 0 6 0 】

HVダンプ対象領域読出処理部の呼び出し部 4 1 0 - m は、HVダンプ対象領域読出処理部 3 6 0 を呼び出す。

カーネルダンプ対象メモリ読出処理部 4 1 1 - m は、カーネルダンプ対象領域（パニック時にOS 4 0 1 - m のカーネルが使用していたメモリ領域）のメモリ内容を読み出す。

【 0 0 6 1 】

HVダンプ採取処理部 4 1 2 - m は、HVダンプ対象領域読出処理部 3 6 0 からHVダンプ対象領域読出処理部 3 6 0 が読み出したメモリ内容を受信し、ダンプファイルを生成する。

カーネルダンプ採取処理部 4 1 3 - m は、読み出したカーネルダンプ対象領域のメモリ内容をファイルに保存（ダンプファイルを生成）する。

【 0 0 6 2 】

パニック処理部 4 1 4 - m は、ドメイン 2 0 1 - m を緊急停止（パニック）させる。

カーネルダンプ対象領域通知処理部 4 1 5 - m は、パニック時にOS 4 0 1 - m のカーネルが使用しているメモリ領域をハイパーバイザ 3 5 1 に通知する。

【 0 0 6 3 】

メモリDR組み込み処理部 4 1 6 - m は、ドメイン 2 0 1 - m にメモリ領域を組み込む。

メモリDR切り離し処理部 4 1 7 - m は、ドメイン 2 0 1 - m からメモリ領域を切り離す。

【 0 0 6 4 】

空きメモリ追加処理部 4 1 8 - m は、メモリ管理部 4 0 2 - m にダンプ済みのメモリ領域を通知する。

ダンプ専用ドメイン停止処理部 4 1 9 - m は、ダンプを採取した後にダンプを採取するドメイン（ダンプ専用ドメイン）を停止する。

【 0 0 6 5 】

カーネルダンプフラグリセット処理部 4 2 0 - m は、ハイパーバイザ 3 5 1 にカーネルダンプフラグのリセットを指示する。

カーネルダンプフラグ通知部 4 2 1 - m は、ハイパーバイザ 3 5 1 にカーネルダンプフラグを通知する。カーネルダンプフラグ通知部 4 2 1 - m は、カーネルのメモリダンプを実行する必要がある場合にカーネルダンプフラグをハイパーバイザ 3 5 1 に通知する。例えば、カーネルダンプフラグ通知部 4 2 1 - m は、カーネルのメモリダンプを採取しない場合はカーネルダンプフラグの値を“ 0 : カーネルのメモリダンプを採取しない”、ダンプ採取用ドメインでカーネルのメモリダンプを採取する場合にカーネルダンプフラグの値を“ 1 : ダンプ採取用ドメインで採取”、メモリDynamic Reconfiguration機能を用いてカーネルのメモリダンプを採取する場合にカーネルダンプフラグの値を“ 2 : メモリDynamic Reconfiguration機能を用いて採取”とする。

【 0 0 6 6 】

（第 1 の実施の形態）

第 1 の実施の形態では、制御ドメインを用いてハイパーバイザのメモリダンプを採取する。

【 0 0 6 7 】

図 7 A、7 B は、第 1 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

初期状態において、ドメイン 2 0 1 - 1 ~ 2 0 1 - 3 およびOS 4 0 1 - 1 ~ 4 0 1 - 3 は、起動され運用状態となっており、ドメイン 2 0 1 - 4 およびOS 4 0 1 - 4 は起動されていないものとする。

【 0 0 6 8 】

ステップ S 5 0 1 において、ハイパーバイザ 3 5 1 は、致命的なエラーを検出する。

ステップ S 5 0 2 において、ドメイン緊急停止指示部 3 5 2 は、運用状態の論理ドメイ

10

20

30

40

50

ン、すなわち制御ドメイン 201 - 1 およびドメイン 201 - 2、201 - 3 に緊急停止を指示する。

【0069】

ステップ S 5 0 3 において、OS 4 0 1 - i (i = 1 ~ 3) は、緊急停止指示を受信し、OS 4 0 1 - i を緊急停止させる。

ステップ S 5 0 4 において、HVダンプ対象領域通知処理部 3 5 4 は、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 からHVダンプ対象領域情報を読み出し、ファームウェア 3 1 1 に通知する。HVダンプ対象領域情報は、ハイパーバイザ 3 5 1 が使用しているメモリ領域(ダンプ対象領域)を示す情報であり、メモリ領域の開始アドレス(PA Base)およびサイズの情報を含む。HVダンプ対象領域情報は、図 8 に示すような形式であり、ブロックの番号、ブロックの物理メモリの開始アドレス(PA Base)、およびブロックのサイズが対応付けられている。また、HV再起動処理部 3 5 7 は、ハイパーバイザ 3 5 1 を停止する(HVアポート)。

【0070】

ステップ S 5 0 6 において、ダンプ対象領域情報/HVダンプフラグ格納処理部 3 1 2 は、HVダンプ対象領域情報を受信する。

ステップ S 5 0 7 において、ダンプ対象領域情報/HVダンプフラグ格納処理部 3 1 2 は、受信したHVダンプ対象領域情報をダンプ対象領域情報として、ダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3 に格納する。また、HVダンプフラグ設定部 3 1 4 は、HVダンプフラグをTUREに設定し、ダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3 に格納する。

【0071】

ステップ S 5 0 8 において、ファームウェア 3 1 1 は、メモリの内容を保持したまま、物理パーティションの再起動処理を開始する。

ステップ S 5 0 9 において、メモリ初期化処理部 3 1 5 は、メモリの初期化処理を開始する。まず、例えば、メモリの先頭のアドレスを初期化処理対象の領域として設定する。

【0072】

ステップ S 5 1 0 において、メモリ初期化処理部 3 1 5 は、ダンプ対象領域情報を参照し、初期化処理対象の領域がダンプ対象領域情報で指定される領域、すなわちダンプ対象領域であるか否か判定する。初期化処理対象の領域がダンプ対象領域である場合、初期化処理対象の領域の内容を保持したまま、制御はステップ S 5 1 2 に進み、ダンプ対象領域でない場合、制御はステップ S 5 1 1 に進む。

【0073】

ステップ S 5 1 1 において、メモリ初期化処理部 3 1 5 は、初期化処理対象の領域を初期化する。

ステップ S 5 1 2 において、メモリ初期化処理部 3 1 5 は、ダンプ対象領域以外のすべての領域に対する初期化処理を行ったか判定する。ダンプ対象領域以外のすべての領域に対する初期化処理を行った場合、制御はステップ S 5 1 3 に進み、ダンプ対象領域以外のすべての領域に対する初期化処理を行っていない場合、未処理の領域(例えば、ダンプ対象領域であるかチェック済みの領域の次のアドレス)を初期化処理対象の領域とし、制御はステップ S 5 1 0 に戻る。

【0074】

ステップ S 5 1 3 において、HV使用領域変更部 3 1 6 は、ハイパーバイザ 3 5 1 が使用する領域をダンプ対象領域情報で示される領域以外の領域に変更する。なお、初期化処理対象の領域として、少なくともハイパーバイザ 3 5 1 が使用する変更後の領域を初期化対象としてもよい。

【0075】

ステップ S 5 1 4 において、HV再起動命令部 3 1 7 は、ハイパーバイザ 3 5 1 に再起動を指示する。ダンプ対象領域情報/HVダンプフラグ通知部 3 1 8 は、ダンプ対象領域情報/HVダンプフラグ格納領域 3 1 3 からダンプ対象領域情報およびHVダンプフラグを読み出し

、ハイパーバイザ 3 5 1 に通知する。

【 0 0 7 6 】

ステップ S 5 1 5 において、HVダンプ対象領域情報/HVダンプフラグ格納処理部 3 5 5 は、ダンプ対象領域情報およびHVダンプフラグを受信し、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 に格納する。尚、HVダンプ対象領域情報/HVダンプフラグ格納処理部 3 5 5 は、受信したダンプ対象領域情報をHVダンプ対象領域情報として格納する。

【 0 0 7 7 】

ステップ S 5 1 6 において、HV再起動処理部 3 5 7 は、ハイパーバイザ 3 5 1 を再起動する。ただし、HVダンプ対象領域情報で指定されるメモリ領域は使用しない。

ステップ S 5 1 7 において、OS再起動命令部 3 5 8 は、OS 4 0 1 - 1 ~ 4 0 1 - 3 に再起動を指示する。

【 0 0 7 8 】

ステップ S 5 1 8 において、OS再起動処理部 4 0 9 - 2、4 0 9 - 3 は、それぞれOS 4 0 1 - 2、4 0 1 - 3 を再起動する。

ステップ S 5 1 9 において、OS 4 0 1 - 2、4 0 1 - 3 は、業務を再開する。

【 0 0 7 9 】

ステップ S 5 2 0 において、OS 4 0 1 - 2、4 0 1 - 3 は、通常の運用状態となる。

ステップ S 5 2 1 において、OS再起動処理部 4 0 9 - 1 は、OS 4 0 1 - 1 を再起動する。

【 0 0 8 0 】

ステップ S 5 2 2 において、OS 4 0 1 - 1 は、業務を再開する。

ステップ S 5 2 3 において、HVメモリダンプ判断部 4 0 8 - 1 は、ハイパーバイザ 3 5 1 にHVダンプフラグの送信を要求する。

【 0 0 8 1 】

ステップ S 5 2 4 において、HVメモリダンプフラグ読出・送信部 3 5 9 は、要求を受信すると、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 からHVダンプフラグを読み出し、OS 4 0 1 - 1 に送信する。

【 0 0 8 2 】

ステップ S 5 2 5 において、HVメモリダンプ判断部 4 0 8 - 1 は、HVダンプフラグを受信し、HVダンプフラグがTRUEであるか否か判定する。HVダンプフラグがTRUEの場合、制御はステップ S 5 2 7 に進み、FALSEの場合、制御はステップ S 5 3 1 に進む。

【 0 0 8 3 】

ステップ S 5 2 6 において、HVダンプ対象領域読出処理部の呼び出し部 4 1 0 - 1 は、HVダンプ対象領域読出処理部 3 6 0 を呼び出す。

ステップ S 5 2 7 において、HVダンプ対象領域読出処理部 3 6 0 は、HVダンプ対象領域情報で示されるメモリ領域の内容を読み出し、制御ドメインに送信する。

【 0 0 8 4 】

ステップ S 5 2 8 において、HVダンプ採取処理部 4 1 2 - 1 は、HVダンプ対象領域読出処理部 3 6 0 からHVダンプ対象領域読出処理部 3 6 0 が読み出したメモリ内容を受信し、受信したメモリ内容をファイルに書き出してダンプファイルを生成する。以下、ステップ S 5 2 9、S 5 3 0 とステップ S 5 3 1 の処理が並列に実行される。

【 0 0 8 5 】

ステップ S 5 2 9 において、メモリ開放処理部 3 6 2 は、HVダンプ対象領域情報で指定されるメモリ領域を開放する。また、HVダンプフラグリセット処理部 3 6 3 は、HVダンプフラグをリセット、すなわちFALSEに設定する。HVダンプフラグ通知部 3 6 4 は、ファームウェア 3 1 1 にHVダンプフラグを通知する。

【 0 0 8 6 】

ステップ S 5 3 0 において、メモリ開放処理部 3 2 0 は、ダンプ対象領域情報をクリアする。また、HVダンプフラグリセット処理部 3 2 1 は、HVダンプフラグをリセット、すなわちFALSEに設定する。

10

20

30

40

50

【 0 0 8 7 】

ステップ S 5 3 1 において、OS 4 0 1 - 1 は、通常の運用状態となる。

第 1 の実施の形態に係るメモリダンプ生成処理によれば、エラーを検出してハイパーバイザおよびオペレーティングシステムを再起動する場合、メモリダンプのサイズが大きい場合でも別のメモリ等にコピーを行っていないので、速やかにハイパーバイザおよびオペレーティングシステムを再起動できる。これにより、業務停止時間を短縮することができる。

【 0 0 8 8 】

ここで、第 1 の実施の形態に係るメモリダンプ生成処理の変形例について説明する。

変形例では、稼働中のハイパーバイザのメモリダンプの採取（ハイパーバイザのライブダンプと呼ぶ）が行われる。

10

【 0 0 8 9 】

図 7 C は、第 1 の実施の形態に係るメモリダンプ生成処理の変形例のフローチャートである。

変形例のフローチャートは、図 7 A、7 B の第 1 の実施の形態に係るメモリダンプ生成処理のフローチャートにステップ S 5 3 2、S 5 3 3 が追加され、ステップ S 5 2 5 において NO と判定された場合に、制御がステップ S 5 3 2 に進むものである。

【 0 0 9 0 】

図 7 C では、図 7 A、7 B に対する変更箇所について記載し、その他の部分については同様であるため記載は省略されている。

20

変形例において、例えば、HVダンプフラグのデータ構造を 0：採取せず、1：異常時の HVダンプ、2：HVライブダンプのように変更することができる。HVメモリダンプ判断部 4 0 8 - 1 は、HVダンプフラグが 1 の場合、HVダンプフラグが TRUE と判定し、HVダンプフラグが 0 または 2 の場合、HVダンプフラグが FALSE と判定する。また、HVメモリダンプ判断部 4 0 8 - 1 は、HVダンプフラグが 2 の場合、HVダンプライブフラグが TRUE と判定する。

【 0 0 9 1 】

ステップ S 5 3 2 において、HVメモリダンプ判断部 4 0 8 - 1 は、HVライブダンプフラグが TRUE であるか否か判定する。HVダンプフラグが TRUE の場合（すなわち、HVダンプフラグが 2 の場合）、制御はステップ S 5 3 3 に進み、FALSE の場合、制御はステップ S 5 3 1 に進む。

30

【 0 0 9 2 】

ステップ S 5 3 3 において、HVライブダンプ処理が行われる。詳細には、HVダンプ対象領域読出処理部の呼び出し部 4 1 0 - 1 は、HVダンプ対象領域読出処理部 3 6 0 を呼び出す。HVダンプ対象領域読出処理部 3 6 0 は、稼働中のハイパーバイザ 3 5 1 が使用しているメモリ領域の内容を読み出し、制御ドメインに送信する。HVダンプ採取処理部 4 1 2 - 1 は、HVダンプ対象領域読出処理部 3 6 0 が読み出したメモリ内容を受信し、受信したメモリ内容をファイルに書き出してハイパーバイザのダンプファイルを生成する。

【 0 0 9 3 】

上記のように、稼働中のハイパーバイザのメモリダンプの採取では、ハイパーバイザを停止・再起動しないまま、ハイパーバイザが使用するメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出している。

40

【 0 0 9 4 】

（第 2 の実施の形態）

第 2 の実施の形態では、ハイパーバイザのメモリダンプに加えて、OS のカーネルのメモリダンプを行う。

【 0 0 9 5 】

図 9 A、9 B、9 C は、第 2 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

初期状態において、ドメイン 2 0 1 - 1 ~ 2 0 1 - 3 および OS 4 0 1 - 1 ~ 4 0 1 - 3 は、起動され運用状態となっており、ドメイン 2 0 1 - 4 および OS 4 0 1 - 4 は起動

50

されていないものとする。

【0096】

ステップS601において、ハイパーバイザ351に致命的なエラーが発生する。

ステップS602において、ハイパーバイザ351は、致命的なエラーを検出する。

ステップS603において、割り込み処理部370は、運用状態のOS、すなわちOS401-i (i = 1 ~ 3) に割り込み処理を通知し、OSパニック指示部353は、OS401-i にパニックを指示する。

【0097】

ステップS604において、パニック処理部414-iは、パニック指示を受信し、OS401-iをパニックさせる。

ステップS605において、カーネルダンプ対象領域通知処理部415-iは、ハイパーバイザ351にカーネルダンプ対象領域情報を通知する。カーネルダンプ対象領域情報は、OS401-iのカーネルが使用しているメモリ領域(ダンプ対象領域)を示す情報であり、メモリ領域の開始アドレス(RA Base)およびサイズの情報を含む。カーネルダンプ対象領域情報は、図10に示すような形式であり、ブロックの番号、ブロックのメモリの開始アドレス(RA Base)、およびブロックのサイズが対応付けられている。

【0098】

尚、ステップS604およびS605は、パニック指示を受信した論理ドメインごとにそれぞれ実行される。

ステップS606において、PA-RAマッピング処理部368は、通知された開始アドレス(RA Base)をRA BaseからPA Baseの開始アドレス(PA Base)に変換するRA-PA変換を行う。

【0099】

ステップS607において、HVダンプ対象領域通知処理部354は、ハイパーバイザ351が使用しているメモリ領域を示すHVダンプ対象領域情報をファームウェア311に通知する。さらに、HVダンプ対象領域通知処理部354は、OS401-iから受信したカーネルダンプ対象領域情報をファームウェア311に通知する。尚、通知されるカーネルダンプ対象領域情報は、RA BaseからPA Baseに変換された開始アドレス(PA Base)およびサイズを含む。実施の形態では、停止した論理ドメインに対応する3個のカーネルダンプ対象領域情報が通知される。

【0100】

ステップS608において、ダンプ対象領域情報/HVダンプフラグ格納処理部312は、受信したHVダンプ対象領域情報および受信したカーネルダンプ対象領域情報をダンプ対象領域情報として、ダンプ対象領域情報/HVダンプフラグ格納領域313に格納する。また、HVダンプフラグ設定部314は、HVダンプフラグをTUREに設定し、ダンプ対象領域情報/HVダンプフラグ格納領域313に格納する。

【0101】

ステップS609において、HV再起動処理部357は、ハイパーバイザ351を停止する(HVアポルト)。

ステップS610において、メモリ初期化処理部315は、ダンプ対象領域情報で示される領域以外のメモリ領域を初期化する。すなわち、メモリ初期化処理部315は、ハイパーバイザ351が使用していた領域とパニック時にOS401-iのカーネルが使用していた領域以外のメモリ領域を初期化する。

【0102】

ステップS611において、HV使用領域変更部316は、HV使用領域変更部316は、ハイパーバイザ351が使用する領域をダンプ対象領域情報で示される領域以外の領域に変更する。HV再起動命令部317は、ハイパーバイザ351に再起動を指示する。ダンプ対象領域情報/HVダンプフラグ通知部318は、ダンプ対象領域情報/HVダンプフラグ格納領域313からダンプ対象領域情報およびHVダンプフラグを読み出し、ハイパーバイザ351に通知する。ダンプ対象領域情報には、HVダンプ対象領域情報およびカーネルダンプ

10

20

30

40

50

対象領域情報が含まれている。HVダンプ対象領域情報/HVダンプフラグ格納処理部 3 5 5 は、ダンプ対象領域情報の内のHVダンプ対象領域情報およびHVダンプフラグを受信し、HVダンプ対象領域情報/HVダンプフラグ格納領域 3 5 6 に格納する。カーネルダンプ対象領域情報/カーネルダンプフラグ格納処理部 3 6 6 は、ダンプ対象領域情報の内のカーネルダンプ対象領域情報を受信し、カーネルダンプ対象領域情報/カーネルダンプフラグ格納領域 3 6 7 に格納する。

【 0 1 0 3 】

ステップ S 6 1 2 において、HV再起動処理部 3 5 7 は、ハイパーバイザ 3 5 1 を起動する。

ステップ S 6 1 3 において、メモリ初期化処理部 3 7 2 は、カーネルダンプ対象領域情報で示される領域以外のメモリ領域を初期化する。

10

【 0 1 0 4 】

ステップ S 6 1 4 において、PA-RAマッピング処理部 3 6 8、OS再起動命令部 3 5 8、およびダンプ専用ドメイン起動処理部 3 6 5 は、カーネルダンプフラグの値をチェックする。以下、カーネルダンプフラグの値に応じた処理が実行される。例えばPA-RAマッピング処理部 3 6 8 は、カーネルダンプフラグが “ 1 : ダンプ採取用ドメインで採取 ” の場合、パニック発生時に OS 4 0 1 - 1 ~ 4 0 1 - 3 のカーネルが使用していたメモリのPAをダンプ採取用ドメイン 2 0 4 - 4 のRAに割り当てる。

【 0 1 0 5 】

以下、ステップ S 6 2 1、ステップ S 6 2 2 ~ S 6 2 6、およびステップ S 6 3 2 ~ S 6 3 5 の処理が別々に並列して実行される。

20

ただし、カーネルダンプフラグが “ 1 : ダンプ採取用ドメインで採取 ” の場合、ステップ S 6 2 6、S 6 3 5 は実行されず、“ 2 : メモリDynamic Reconfiguration機能を用いて採取 ” の場合、ステップ S 6 2 1 は実行されない。

【 0 1 0 6 】

ここでは、ステップ S 6 2 1 は、ダンプ採取用ドメイン 2 0 4 - 4 に関する処理であり、ステップ S 6 2 2 ~ S 6 2 6 は、制御ドメイン 2 0 4 - 1 に関する処理であり、ステップ S 6 3 2 ~ S 6 3 5 は、論理ドメイン 2 0 4 - 2、2 0 4 - 3 に関する処理である。

【 0 1 0 7 】

ステップ S 6 2 1 において、ダンプ採取用ドメインによるカーネルのメモリダンプ生成処理が行われる。ダンプ採取用ドメインによるカーネルのメモリダンプ生成処理の詳細については後述する。

30

【 0 1 0 8 】

ステップ S 6 2 2 において、PA-RAマッピング処理部 3 6 8 は、ドメイン 2 0 1 - 1 の物理アドレス(PA)とリアルアドレス(RA)間のマッピングを以下の1)、2)のように変更する。それにより、OS 4 0 1 - 1 を再起動してもパニック時のカーネルおよびパニック時のハイパーバイザ 3 5 1 が使用していたメモリ領域のデータは上書きされなくなる。

1)パニック発生時にカーネルおよびハイパーバイザが使用していたメモリの物理アドレスは、再起動するドメインのリアルアドレスに割り当てないようにする。かつ、

2)再起動前後で、該当ドメインが使用できるメモリサイズがなるべく変化しないようにする。

40

【 0 1 0 9 】

ただし、再起動するドメインに割り当て可能な物理メモリが所定値より不足する場合は、1)を優先する。

尚、パニック発生時にどの領域をカーネルおよびハイパーバイザ 3 5 1 が利用していたかは、HVダンプ対象領域情報およびカーネルダンプ対象領域情報を参照することにより判断される。

【 0 1 1 0 】

ステップ S 6 2 3 において、OS再起動命令部 3 5 8 は、OS 4 0 1 - 1 に再起動を指示する。また、OS再起動命令部 3 5 8 は、カーネルダンプフラグが “ 2 : メモリDynamic Reco

50

nfiguration機能を用いて採取”の場合、メモリDR機能を用いたカーネルのメモリダンプ生成処理を行う旨を再起動指示に含める。指示を受信したOS起動処理部409-1は、OS401-1を起動する。

【0111】

ステップS624において、OS401-1は、業務を再開する。

ステップS625において、メモリDynamic Reconfiguration (DR) 機能を用いたカーネルのメモリダンプ生成処理が行われる。メモリDR機能を用いたカーネルのメモリダンプ生成処理の詳細については後述する。ステップS626において、ハイパーバイザのメモリダンプ生成処理が行われる。ステップS626は、図7BのステップS523~S531の処理と同様であるため説明は省略する。

【0112】

ステップS632において、PA-RAマッピング処理部368は、ドメイン201-2, 201-3の物理アドレス(PA)とリアルアドレス(RA)間のマッピングを以下の1)、2)のように変更する。それにより、OS401-1を再起動してもパニック時のカーネルおよびパニック時のハイパーバイザ351が使用していたメモリ領域のデータは上書きされなくなる。

1)パニック発生時にカーネルおよびハイパーバイザが使用していたメモリの物理アドレスは、再起動するドメインのリアルアドレスに割り当てないようにする。かつ、

2)再起動前後で、該当ドメインが使用できるメモリサイズがなるべく変化しないようにする。

【0113】

ただし、再起動するドメインに割り当て可能な物理メモリが所定値より不足する場合は、1)を優先する。

尚、パニック発生時にどの領域をカーネルおよびハイパーバイザ351が利用していたかは、HVダンプ対象領域情報およびカーネルダンプ対象領域情報を参照することにより判断される。

【0114】

ステップS633において、OS再起動命令部358は、OS401-2、401-3に再起動を指示する。また、OS再起動命令部358は、カーネルダンプフラグが“2:メモリDynamic Reconfiguration機能を用いて採取”の場合、メモリDR機能を用いたカーネルのメモリダンプ生成処理を行う旨を再起動指示に含める。指示を受信したOS起動処理部409-2、409-3は、OS401-2, 401-3をそれぞれ起動する。

【0115】

ステップS634において、OS401-2, 401-3は、それぞれ業務を再開する。

ステップS635において、メモリDR機能を用いたカーネルのメモリダンプ生成処理が行われる。

【0116】

以下、カーネルのメモリダンプ生成処理の詳細について説明する。

カーネルのメモリダンプ生成処理は、(1)ダンプ採取用ドメインによるメモリダンプを採取する方法(ステップS621)、または(2)メモリDynamic Reconfiguration機能を用いてメモリダンプを採取する方法(ステップS626、S635)のいずれかが用いられる。

【0117】

(1)ダンプ採取用ドメインでメモリダンプを採取する方法

ダンプ採取用ドメイン201-4は、複数のドメイン201が存在するシステムでも、それぞれの論理ドメイン毎に用意する必要はなく、システムで1つあれば良い。ダンプ採取用ドメイン201-4が1つの場合、複数の論理ドメイン201で同時にパニックが発生した場合は1ドメインずつメモリダンプを採取することになるが、ダンプ採取が完了しているかどうかにかかわらず、速やかに業務が再開できるため、業務への影響はない。

【0118】

10

20

30

40

50

ダンプ採取用ドメイン 201 - 4 では、パニックが発生した論理ドメインの業務を引き継ぐ必要はないため、メモリダンプを採取するために必要となる下記のハードウェア資源があれば良い。

- ・パニックが発生した論理ドメインのOSのカーネルがパニック時に使用していた物理メモリ領域
- ・1個以上のCPU
- ・ダンプファイルを格納するディスクとディスクを使用するために必要なI/O資源

【0119】

図11は、ダンプ採取用ドメインによるカーネルのメモリダンプ生成処理のフローチャートである。

10

図11は、図9BのステップS621に対応する。

【0120】

ここでは、OS 401 - i のカーネルのメモリダンプ生成処理について説明する。

ステップS651において、ダンプ専用ドメイン起動処理部365は、ダンプ採取用ドメイン201 - 4をファームウェアモードで起動する。ファームウェアモードとは、OSを起動しないモード、すなわちOSを起動する前に停止するモードである。OSを起動しないことにより、ダンプ対象領域が書き換えられてしまうことを防ぐ。

【0121】

ステップS652において、カーネルダンプ対象メモリ読出処理部411 - 4は、パニック発生時にオペレーティングシステム401 - iのカーネルが使用していたメモリ領域（カーネルダンプ対象領域）を読み出す。尚、カーネルダンプ対象領域の情報（開始アドレス（RA Base）やサイズ等）は、ファームウェア311またはハイパーバイザ351からの通知により得る。

20

【0122】

ステップS653において、カーネルダンプ採取処理部413 - 4は、読み出したメモリ内容をファイルに書き出してダンプファイルを生成する。

ステップS654において、ダンプ専用ドメイン停止処理部419 - 4は、ダンプ採取ドメイン201 - 4を停止する。そして、ダンプ専用ドメイン停止処理部419 - 4は、カーネルダンプ対象領域を使用可能な未使用のメモリ、すなわち空きメモリとするようにハイパーバイザ351のメモリ管理部361へ通知する。また、カーネルダンプフラグリセット処理部420 - 4は、ハイパーバイザ351にカーネルダンプフラグのリセットを指示する。リセット指示を受信したカーネルダンプフラグリセット処理部373は、カーネルダンプフラグをリセットする。

30

【0123】

ステップS655において、メモリ管理部361は、カーネルダンプ対象領域を他の論理ドメイン201 - iからも使用可能な空きメモリとする。

【0124】

図12は、ダンプ採取用ドメインによるメモリダンプの採取を示す図である。

図12の左側は運用状態（およびパニック時）、真ん中は再起動時、右側はダンプ採取用ドメインによるメモリダンプの採取時を示す。

40

ここでは、論理ドメイン201 - 1の処理について記載している。尚、ドメイン201 - 2、201 - 3においても同様の処理が実行されるので、詳細は省略する。

【0125】

図12の左側の運用状態において、PAのある領域が論理ドメイン201 - 1のRAのある領域にマッピングされている。

OS 401 - 1のパニック時にOS 401 - 1のカーネルが使用していた領域はカーネルダンプ対象領域となる。

【0126】

OS 401 - 1のパニック後、PA-RAマッピングの変更が行われ（ステップS622）、論理ドメイン201 - 1には、パニック時にOS 401 - 1のカーネルが使用していた領域

50

(カーネルダンプ対象領域)とは異なるPAの領域が割り当てられ、OS 4 0 1 - 1は再起動する(図12の真ん中)。

【0127】

図12の右側のダンプ時において、ダンプ採取専用ドメイン201-4のRAには、パニック時にOS 4 0 1 - 1のカーネルが使用していたPAの領域(カーネルダンプ対象領域)が割り当てられる。ダンプ採取専用ドメイン201-4は、カーネルダンプ対象領域を読み出して、ダンプファイルを生成する。

【0128】

図13は、ダンプ採取用ドメインによるメモリダンプの採取におけるPA-RAマッピング情報を示す図である。

図13の左側は運用状態(およびパニック時)、真ん中はダンプ時、右側はダンプ後を示す。

【0129】

ここでは、ドメイン201-1(制御ドメイン#0)とダンプ採取専用ドメイン201-4(ダンプ採取専用ドメイン#3)のPA-RAマッピングについて記載している。

PA-RAマッピング情報は、ドメイン、開始アドレス(PA Base)、サイズ、および開始アドレス(RA Base)が対応付けられて記載されている。

【0130】

図13の左側のパニック時において、開始アドレス(PA Base)がxxxxx、サイズが8GBである領域が制御ドメイン#0の開始アドレス(RA Base)がaaaaaである領域にマッピングされている(図12の左側に対応)。この領域がカーネルダンプ対象領域となる。

【0131】

OS 4 0 1 - 1のパニック後、PA-RAマッピングの変更が行われ(ステップS622)、PA-RAマッピング情報は図13の真ん中に示すようになる。

図13の真ん中のダンプ時において、開始アドレス(PA Base)がxxxxx、サイズが8GBである領域がダンプ採取専用ドメイン#3の開始アドレス(RA Base)がaaaaaである領域にマッピングされている。すなわち、パニック時の制御ドメイン#0のPAの領域がダンプ採取専用ドメイン#3のRAにマッピングされている。また、開始アドレス(PA Base)がyyyyy、サイズが8GBである領域が制御ドメイン#0の開始アドレス(PA Base)がaaaaaである領域にマッピングされている。すなわち、新たなPAの領域が再起動後の制御ドメイン#0に割り当てられている(図12の右側に対応)。

【0132】

ダンプファイルの生成後、カーネルダンプ対象領域は他のドメインからも使用可能な空きメモリとなる(ステップS655)。

すなわち、図13の右側のダンプ後において、ダンプ採取専用ドメイン#3のマッピング情報は削除される。

【0133】

ダンプ採取用ドメインでメモリダンプを採取する方法によれば、異常を検出したドメインではなく、別のドメインでダンプを採取するため、ダンプ採取処理中に再度異常を検出してハングアップする等の二次被害が発生する可能性が低くなる。

【0134】

ダンプ採取用ドメインでメモリダンプを採取する方法によれば、Capacity on Demand (CoD)のような、ユーザが使用したハードウェア資源(CPU、メモリ、ディスク等)の量や時間に応じて課金を行うシステムにおいて、ダンプ採取のために使用するハードウェア資源に対する課金を行わないようにすることが容易に実現でき、料金の適正化を図ることができる。

【0135】

(2)メモリDynamic Reconfiguration機能を用いてメモリダンプを採取する方法

ここでは、論理ドメイン201-1の処理(ステップS625)について説明する。尚、論理ドメイン201-2、201-3の処理(ステップS635)も同様の処理が実行

10

20

30

40

50

されるので、詳細は省略する。

【0136】

図14は、メモリDynamic Reconfiguration機能を用いたカーネルのメモリダンプ生成処理のフローチャートである。

図14は、図9CのステップS625に対応する。

【0137】

ステップS641において、メモリDR組み込み処理部416-1は、メモリのDynamic Reconfiguration機能を使用して、パニック発生時にOS401-1のカーネルが使用していたメモリ領域(カーネルダンプ対象領域)をドメイン201-1に組み込む。なお、カーネルダンプ対象領域の情報(開始アドレス(RA Base)やサイズ等)は、ファームウェア311またはハイパーバイザ351からの通知により得る。

10

【0138】

ステップS642において、カーネルダンプ対象メモリ読出処理部411-1は、組み込んだメモリ領域を読み出す。

ステップS643において、カーネルダンプ採取処理部413-1は、読み出したメモリ内容をファイルに書き出してダンプファイルを生成する。

【0139】

ステップS644において、メモリDR切り離し処理部417-1は、メモリのDynamic Reconfiguration機能を使用して、パニック発生時にOS401-1のカーネルが使用していたメモリ領域をドメイン201-1から切り離して、切り離した領域を空きメモリとするようにメモリ管理部361に通知する。また、カーネルダンプフラグリセット処理部420-1は、ハイパーバイザ351にカーネルダンプフラグのリセットを指示する。リセット指示を受信したカーネルダンプフラグリセット処理部373は、カーネルダンプフラグをリセットする。

20

【0140】

ステップS645において、メモリ管理部361は、切り離した領域を他のドメイン201-2、201-3からも使用可能な空きメモリとする。

また、ステップS644およびS645の代わりに、空きメモリ追加処理部418-1は、パニック時にOS401-1のカーネルが使用していたメモリ領域(すなわち、ダンプ済み領域)を使用可能な未使用のメモリ、すなわち空きメモリとするようにメモリ管理部402-1へ通知し、メモリ管理部402-1はダンプ済み領域を空きメモリとする処理を行っても良い。

30

【0141】

図15は、メモリDynamic Reconfiguration機能を用いたメモリダンプの採取を示す図である。

図15の左側は運用状態(およびパニック時)、真ん中は再起動時、右側はダンプ採用ドメインによるメモリダンプの採取時を示す。

ここでは、ドメイン201-1の処理について記載している。尚、ドメイン201-2、201-3においても同様の処理が実行されるので、詳細は省略する。

【0142】

図15の左側の運用状態において、PAのある領域がドメイン201-1のRAのある領域にマッピングされている。

40

【0143】

OS401-1のパニック時にOS401-1のカーネルが使用していた領域はカーネルダンプ対象領域となる。

OS401-1のパニック後、PA-RAマッピングの変更が行われ(ステップS622)、ドメイン201-1のRAには、パニック時にOS401-1のカーネルが使用していた領域(カーネルダンプ対象領域)とは異なるPAの領域が割り当てられ、OS401-1は再起動する(図15の真ん中)。

【0144】

50

図15の右側の再起動後のダンプ時において、ドメイン201-1のRAには、パニック時にOS401-1のカーネルが使用していた領域（カーネルダンプ対象領域）が組み込まれる。ドメイン201-1は、カーネルダンプ対象領域を読み出して、ダンプファイルを生成する。

【0145】

図16は、メモリDynamic Reconfiguration機能を用いたメモリダンプの採取におけるPA-RAマッピング情報を示す図である。

図16の左側は運用状態（およびパニック時）、真ん中はダンプ時、右側はダンプ後を示す。

【0146】

ここでは、ドメイン201-1（制御ドメイン#0）のPA-RAマッピングについて記載している。

PA-RAマッピング情報は、ドメイン、開始アドレス（PA Base）、サイズ、および開始アドレス（RA Base）が対応付けられて記載されている。

【0147】

図16の左側のパニック時において、開始アドレス（PA Base）がxxxxx、サイズが8GBである領域が制御ドメイン#0の開始アドレス（RA Base）がaaaaaである領域にマッピングされている（図15の左側に対応）。この領域がカーネルダンプ対象領域となる。

【0148】

OS401-1のパニック後、PA-RAマッピングの変更が行われ、さらにカーネルダンプ対象領域が制御ドメイン#0に組み込まれ、PA-RAマッピング情報は図16の真ん中に示すようになる。

【0149】

図16の真ん中のダンプ時において、開始アドレス（PA Base）がyyyyy、サイズが8GBである領域が制御ドメイン#0のRAの開始アドレス（RA Base）がaaaaaである領域にマッピングされている。さらに、開始アドレス（PA Base）がxxxxx、サイズが8GBである領域が制御ドメイン#0の開始アドレス（RA Base）がbbbbbbである領域にマッピングされている。

【0150】

すなわち、新たなPAの領域が再起動後の制御ドメイン#0に割り当てられ、さらに制御ドメイン#0の再起動後に、カーネルダンプ対象領域が制御ドメイン#0に組み込まれる（図15の右側に対応）。

【0151】

ダンプファイルの生成後、カーネルダンプ対象領域は他のドメインからも使用可能な空きメモリとなる（ステップS645）。

すなわち、図16の右側のダンプ後において、カーネルダンプ対象領域のマッピング情報は削除される。

【0152】

メモリDynamic Reconfiguration機能を用いてメモリダンプを採取する方法によれば、異常を検出したオペレーティングシステムではなく、再起動後の新しいオペレーティングシステムがダンプを採取するため、ダンプ採取処理中に再度異常を検出してハングアップする等の二次被害が発生する可能性が低くなる。

【0153】

第2の実施の形態に係るメモリダンプ生成処理によれば、エラーを検出してハイパーバイザおよびオペレーティングシステムを再起動する場合、メモリダンプのサイズが大きい場合でも別のメモリ等にコピーを行っていないので、速やかにハイパーバイザおよびオペレーティングシステムを再起動できる。これにより、業務停止時間を短縮することができる。

【0154】

第2の実施の形態に係るメモリダンプ生成処理によれば、ハイパーバイザおよびカーネ

10

20

30

40

50

ルのメモリダンプを採取することで、ハイパーバイザおよびドメインの両方に起因したエラーであっても、効果的にエラーの解析を行うことができる。

【0155】

(第3の実施の形態)

第3の実施の形態では、OSでエラーが検出され、カーネルのメモリダンプが行われる。

ここでは、OS 4 0 1 - 1のカーネルのメモリダンプを生成する場合について説明する。

【0156】

図17は、第3の実施の形態に係るメモリダンプ生成処理のフローチャートである。

先ず、メモリ管理部 4 0 2 - 1は、OS 4 0 1 - 1の起動時に、メモリ 2 0 3 - 1の一番小さい(または一番大きい)リアルアドレス(RA)からカーネルが使用するメモリを割り当てる。このように、なるべくカーネルが使用するメモリ領域(ダンプ対象領域)のサイズが小さくなるようにする。また、マッピング情報抽出・格納処理部 4 0 6 - 1は、カーネルが使用しているメモリのダンプを採取/解析するために必要となる情報(例えば、カーネルのテキスト域、データ域、ヒープ域、スタック域など、各セグメントのマッピング情報(論理アドレス、物理アドレス、サイズ等)、アドレス変換テーブル、ページテーブル、各種制御テーブルのマッピング情報)を、マッピング情報格納領域 4 0 7 - 1に書き込む。また、ドメイン 2 0 1 - 1 ~ 2 0 1 - 3およびOS 4 0 1 - 1 ~ 4 0 1 - 3は、起動され運用状態となっており、ドメイン 2 0 1 - 4およびOS 4 0 1 - 4は起動されていないものとする。

【0157】

ステップ S 7 0 1において、OS 4 0 1 - 1に致命的なエラーが発生する。

ステップ S 7 0 2において、OS 4 0 1 - 1は、致命的なエラーを検出する。

ステップ S 7 0 3において、パニック処理部 4 1 4 - 1は、OS 4 0 1 - 1をパニック(緊急停止)させる。

【0158】

ステップ S 7 0 4において、カーネルダンプ対象領域通知処理部 4 1 5 - 1は、緊急停止(パニック)時にOS 4 0 1 - 1のカーネルが使用していたメモリ領域(カーネルダンプ対象領域)の情報(カーネルダンプ対象領域情報)をハイパーバイザ 3 5 1に通知する。また、カーネルダンプフラグ通知部 4 2 1 - 1は、カーネルダンプフラグをハイパーバイザ 3 5 1に通知する。

【0159】

ステップ S 7 0 5において、カーネルダンプ対象領域情報/カーネルダンプフラグ格納処理部 3 6 6は、受信したカーネルダンプ対象領域情報およびカーネルダンプフラグをカーネルダンプ対象領域情報/カーネルダンプフラグ格納領域 3 6 7に格納する。

【0160】

ステップ S 7 0 6において、メモリ初期化部 3 7 2は、カーネルダンプ対象領域情報で示される領域以外のメモリ領域を初期化する。すなわち、メモリ初期化部 3 7 2は、パニック時にOS 4 0 1 - 1のカーネルが使用していたメモリ領域の初期化処理を実施しない(すなわち、データを更新しない)ようにする。それにより、パニック時にOS 4 0 1 - 1のカーネルが使用していたメモリ領域のデータはそのままの状態に残る。

【0161】

ステップ S 7 0 7において、PA-RAマッピング処理部 3 6 8、OS再起動命令部 3 5 8、およびダンプ専用ドメイン起動処理部 3 6 5は、カーネルダンプフラグの値をチェックする。以下、カーネルダンプフラグの値に応じた処理が実行される。例えばPA-RAマッピング処理部 3 6 8は、カーネルダンプフラグが“1:ダンプ採取用ドメインで採取”の場合、パニック発生時にOS 4 0 1 - 1 ~ 4 0 1 - 3のカーネルが使用していたメモリのPAをダンプ採取用ドメイン 2 0 4 - 4のRAに割り当てる。

【0162】

以下、ステップ S 7 0 8とステップ S 7 0 9 ~ S 7 1 2の処理が別々に並列して実行される。

10

20

30

40

50

ただし、カーネルダンプフラグが“ 1 : ダンプ採取用ドメインで採取 ” の場合はステップ S 7 1 2 は実行されず、“ 2 : メモリDynamic Reconfiguration機能を用いて採取 ” の場合はステップ S 7 0 8 は実行されない。

【 0 1 6 3 】

ステップ S 7 0 8 において、ダンプ採取用ドメインによるカーネルのメモリダンプ生成処理が行われる。ステップ S 7 0 8 は、図 9 B のステップ S 6 2 1 の処理と同様であるため説明は省略する。

【 0 1 6 4 】

ステップ S 7 0 9 において、PA-RAマッピング処理部 3 6 8 は、パニックしたドメイン 2 0 1 - 1 の物理アドレス(PA)とリアルアドレス(RA)間のマッピングを以下の1)、2)のように変更する。それにより、OS 4 0 1 - 1 を再起動してもパニック時にOS 4 0 1 - 1 のカーネルが使用していたメモリ領域のデータは上書きされなくなる。

- 1) パニック発生時にカーネルが使用していたメモリの物理アドレスは、再起動するドメインのリアルアドレスに割り当てないようにする。かつ、
- 2) 再起動前後で、該当ドメインが使用できるメモリサイズがなるべく変化しないようにする。

【 0 1 6 5 】

ただし、再起動するドメインに割り当て可能な物理メモリが所定値より不足する場合は、1)を優先する。

ステップ S 7 1 0 において、OS再起動命令部 3 5 8 は、ドメイン 2 0 1 - 1 にOS 4 0 1 - 1 の再起動を指示する。また、OS再起動命令部 3 5 8 は、カーネルダンプフラグが“ 2 : メモリDynamic Reconfiguration機能を用いて採取 ” の場合、メモリDR機能を用いたカーネルのメモリダンプ生成処理を行う旨を再起動指示に含める。OS起動処理部 4 0 9 - 1 は、カーネルが使用していたメモリのダンプをディスク等へ書き出すことなく、OS 4 0 1 - 1 を再起動する。

【 0 1 6 6 】

ステップ S 7 1 1 において、OS 4 0 1 - 1 は、業務を再開する。

ステップ S 7 1 2 において、カーネルのメモリダンプ生成処理が行われる。尚、ステップ S 7 1 2 の処理は、上述のステップ S 6 2 5 の処理と同様であるため説明は省略する。

【 0 1 6 7 】

第 3 の実施の形態に係るメモリダンプ生成処理によれば、エラーを検出してオペレーティングシステムを緊急停止(パニック)する場合、メモリダンプのサイズが大きい場合でも別のメモリ等にコピーを行っていないので、速やかにオペレーティングシステムを再起動できる。これにより、業務停止時間を短縮することができる。

【 0 1 6 8 】

(第 4 の実施の形態)

第 4 の実施の形態では、OSでエラーが検出され、カーネルのメモリダンプが行われ、さらに稼働中のハイパーバイザのメモリダンプの採取(ハイパーバイザのライブダンプと呼ぶ)が行われる。

ここでは、OS 4 0 1 - 1 のカーネルのメモリダンプを生成する場合について説明する。

【 0 1 6 9 】

図 1 8 A、1 8 B は、第 4 の実施の形態に係るメモリダンプ生成処理のフローチャートである。

ステップ S 8 0 1 ~ ステップ S 8 1 1 は、図 1 7 のステップ S 7 0 1 ~ S 7 1 1 とそれぞれ同様の処理であるため、説明は省略する。

【 0 1 7 0 】

以下、ステップ S 8 1 2 とステップ S 8 1 3 は並列に実行される。

ステップ S 8 1 2 において、カーネルのメモリダンプ生成処理が行われる。尚、ステップ S 8 1 2 の処理は、上述のステップ S 6 2 5 の処理と同様であるため説明は省略する。

ステップ S 8 1 3 において、稼働中のハイパーバイザ 3 5 1 のメモリダンプ生成処理が

10

20

30

40

50

制御ドメイン 204 - 1 で行われる。

【0171】

以下、稼動中のハイパーバイザ 351 のメモリダンプ生成処理の詳細について説明する。

図 19 は、稼動中のハイパーバイザのメモリダンプ生成処理のフローチャートである。

【0172】

図 19 は、図 18 B のステップ S 813 に対応する。

第 4 の実施の形態において、例えば、HVダンプフラグのデータ構造を 0 : 採取せず、1 : 異常時の HVダンプ、2 : HVライブダンプのように変更することができる。HVメモリダンプ判断部 408 - 1 は、HVダンプフラグが 0 または 1 の場合、HV のライブダンプを採取しないと判定し、HVダンプフラグが 2 の場合、HV のライブダンプを採取すると判定する。

10

【0173】

ステップ S 831 において、HVメモリダンプ判断部 408 - 1 は、ハイパーバイザ 351 に HVダンプフラグの送信を要求する。

ステップ S 832 において、HVメモリダンプフラグ読出・送信部 359 は、要求を受信すると、HVダンプ対象領域情報/HVダンプフラグ格納領域 356 から HVダンプフラグを読み出し、OS 401 - 1 に送信する。

【0174】

ステップ S 833 において、HVメモリダンプ判断部 408 - 1 は、受信した HVダンプフラグに基づいて、稼動中のハイパーバイザ 351 のライブダンプを採取するか否か判定する。稼動中のハイパーバイザ 351 のライブダンプを採取すると判定された場合、制御はステップ S 834 に進み、採取しないと判定された場合、処理は終了する。

20

【0175】

ステップ S 834 において、HVダンプ対象領域読出処理部の呼び出し部 410 - 1 は、HVダンプ対象領域読出処理部 360 を呼び出す。

ステップ S 835 において、HVダンプ対象領域読出処理部 360 は、現在、ハイパーバイザ 351 が使用しているメモリ領域を読み出し、読み出したメモリ内容を制御ドメイン 204 - 1 に送信する。

【0176】

ステップ S 836 において、HVダンプ採取処理部 412 - 1 は、メモリ内容を受信し、受信したメモリ内容をファイルに書き出してハイパーバイザのダンプファイルを生成する。

30

【0177】

上記のように、稼動中のハイパーバイザのメモリダンプの生成処理では、ハイパーバイザを停止・再起動しないまま、ハイパーバイザが使用するメモリ領域のデータを読み出して、該データをハイパーバイザのダンプファイルとしてファイルに書き出している。

【0178】

第 4 の実施の形態に係るメモリダンプ生成処理によれば、エラーを検出してオペレーティングシステムを再起動する場合、メモリダンプのサイズが大きい場合でも別のメモリ等にコピーを行っていないので、速やかにオペレーティングシステムを再起動できる。これにより、業務停止時間を短縮することができる。

40

【0179】

第 4 の実施の形態に係るメモリダンプ生成処理によれば、ハイパーバイザおよびカーネルのメモリダンプを採取することで、ハイパーバイザおよびドメインの両方に起因したエラーであっても、効果的にエラーの解析を行うことができる。

【0180】

以上、複数の実施の形態を説明してきたが、実施の形態は装置および方法に限らず、プログラムとして構成することも出来るし、該プログラムを格納したコンピュータが読み取り可能な記録媒体として構成することも出来る。記録媒体としては、例えば、フレキシブ

50

ルディスク（FD）、ハードディスクドライブ、光ディスク、光磁気ディスク、CD-ROM、CD-R、DVD-ROM、DVD-RAM、磁気テープ、不揮発性のメモリーカード等が用いられる。

【0181】

例えば、実施の形態のプログラムは、該プログラムを格納した記録媒体から読み出され、メモリ13、23や不揮発性メモリ14に格納される。CPU12、22は、メモリ13、23や不揮発性メモリ14からプログラムを読み出して実行することにより、上述した実施の形態の各種処理を実行する。

【0182】

本発明は、以上に述べた実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲内で種々の構成を取ることができる。例えば、論理ドメインの数は4つの限られるものでなく、任意の数にすることができる。

10

【符号の説明】

【0183】

10	サーバ	
11	システムボード	
12	CPU	
13	メモリ	
14	不揮発性メモリ	
21	サービスプロセッサ	20
22	CPU	
23	メモリ	
31	ディスクユニット	
32	ハードディスクドライブ	
41	通信インタフェース	
51	バス	
61	物理パーティション	
201	論理ドメイン	
202	CPU	
203	メモリ	30
204	ディスク	
311	ファームウェア	
312	ダンプ対象領域情報/HVダンプフラグ格納処理部	
313	ダンプ対象領域情報/HVダンプフラグ格納領域	
314	HVダンプフラグ設定部	
315	メモリ初期化処理部	
316	HV使用領域変更部	
317	HV再起動命令部	
318	ダンプ対象領域情報/HVダンプフラグ通知部	
319	PA-RAマッピング通知部	40
320	メモリ開放処理部	
321	HVダンプフラグリセット処理部	
351	ハイパーバイザ	
352	ドメイン緊急停止指示部	
353	OSパニック指示部	
354	HVダンプ対象領域通知処理部	
355	HVダンプ対象領域情報/HVダンプフラグ格納処理部	
356	HVダンプ対象領域情報/HVダンプフラグ格納領域	
357	HV再起動処理部	
358	OS再起動命令部	50

3 5 9	HVメモリダンプ読出・送信部	
3 6 0	HVダンプ対象メモリ読出処理部	
3 6 1	メモリ管理部	
3 6 2	メモリ開放処理部	
3 6 3	HVダンプフラグリセット処理部	
3 6 4	HVダンプフラグ通知部	
3 6 5	ダンプ専用ドメイン起動処理部	
3 6 6	カーネルダンプ対象領域情報/カーネルダンプフラグ格納処理部	
3 6 7	カーネルダンプ対象領域情報/カーネルダンプフラグ格納領域	
3 6 8	PA-RAマッピング処理部	10
3 6 9	PA-RAマッピング情報格納域	
3 7 0	割り込み処理部	
3 7 1	メモリダンプ処理起動部	
3 7 2	メモリ初期化処理部	
3 7 3	カーネルダンプフラグリセット処理部	
4 0 1	オペレーティングシステム	
4 0 2	メモリ管理部	
4 0 3	ファイル管理部	
4 0 4	プロセス管理部	
4 0 5	割り込み処理部	20
4 0 6	マッピング情報抽出・格納処理部	
4 0 7	マッピング情報格納領域	
4 0 8	HVメモリダンプ判断部	
4 0 9	OS起動処理部	
4 1 0	HVダンプ対象領域読出処理部の呼び出し部	
4 1 1	カーネルダンプ対象メモリ読出処理部	
4 1 2	HVダンプ採取処理部	
4 1 3	カーネルダンプ採取処理部	
4 1 4	パニック処理部	
4 1 5	カーネルダンプ対象領域通知処理部	30
4 1 6	メモリDR組み込み処理部	
4 1 7	メモリDR切り離し処理部	
4 1 8	空きメモリ追加処理部	
4 1 9	ダンプ専用ドメイン停止処理部	
4 2 0	カーネルダンプフラグリセット処理部	
4 2 1	カーネルダンプフラグ通知部	

【図16】

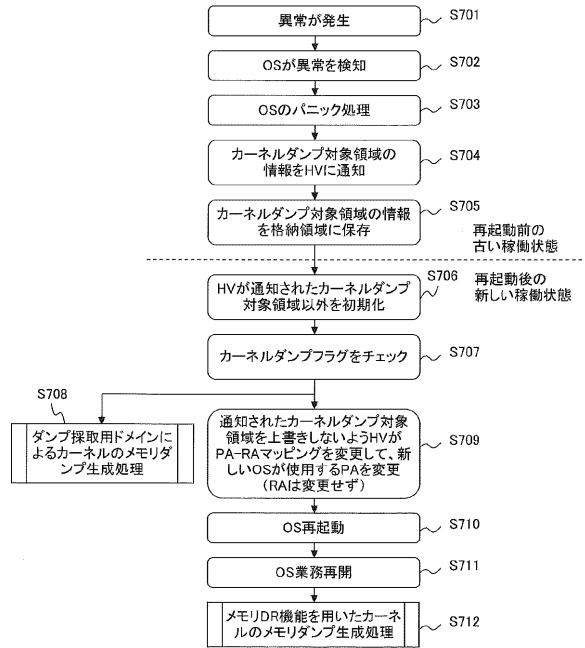
メモリDynamic Reconfiguration機能を用いた
メモリダンプの採取におけるPA-RAマッピング情報を示す図

ダンプ後	ドメイン	#0	yyyyy	8G	aaaaa
	PA Base (開始アドレス)				
	RA Base (開始アドレス)				
ダンプ時	ドメイン	#0	xxxxx	8G	aaaaa
	PA Base (開始アドレス)				
	RA Base (開始アドレス)				bbbbbb
緊急停止時	ドメイン	#0	xxxxx	8G	aaaaa
	PA Base (開始アドレス)				
	RA Base (開始アドレス)				

ダンプ対象領域

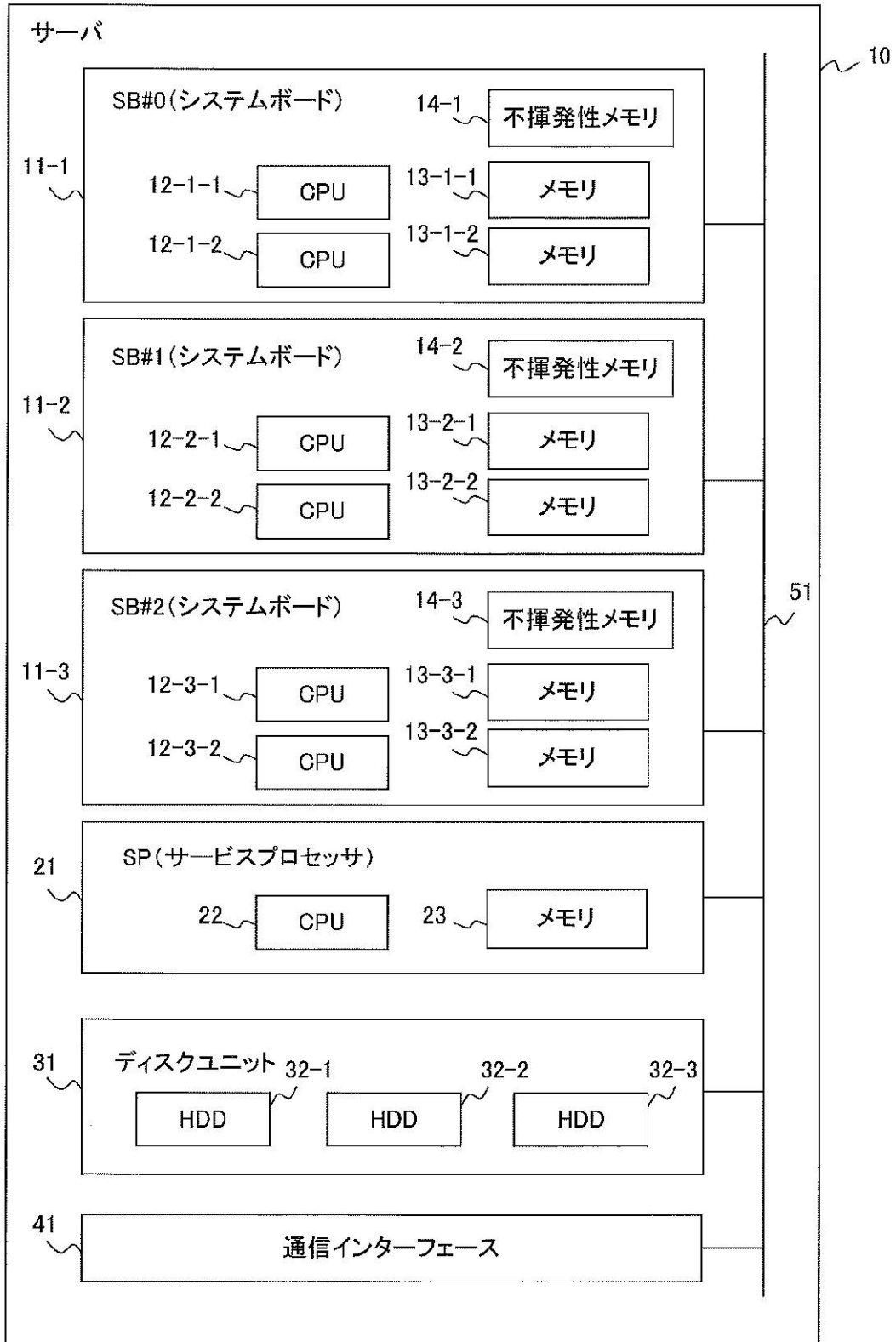
【図17】

第3の実施の形態に係るメモリダンプ生成処理のフローチャート



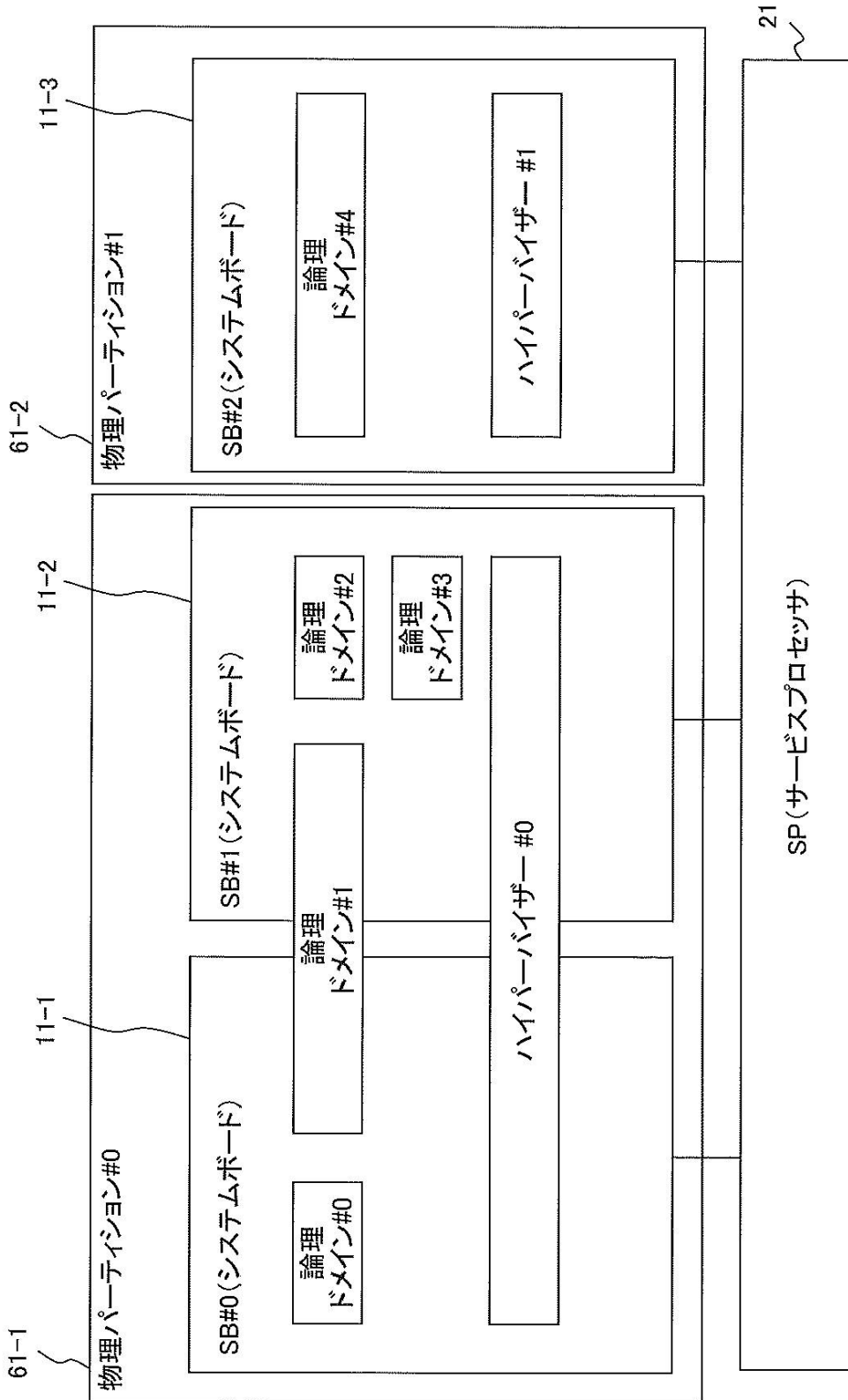
【 図 1 】

実施の形態に係るサーバのハードウェア構成図



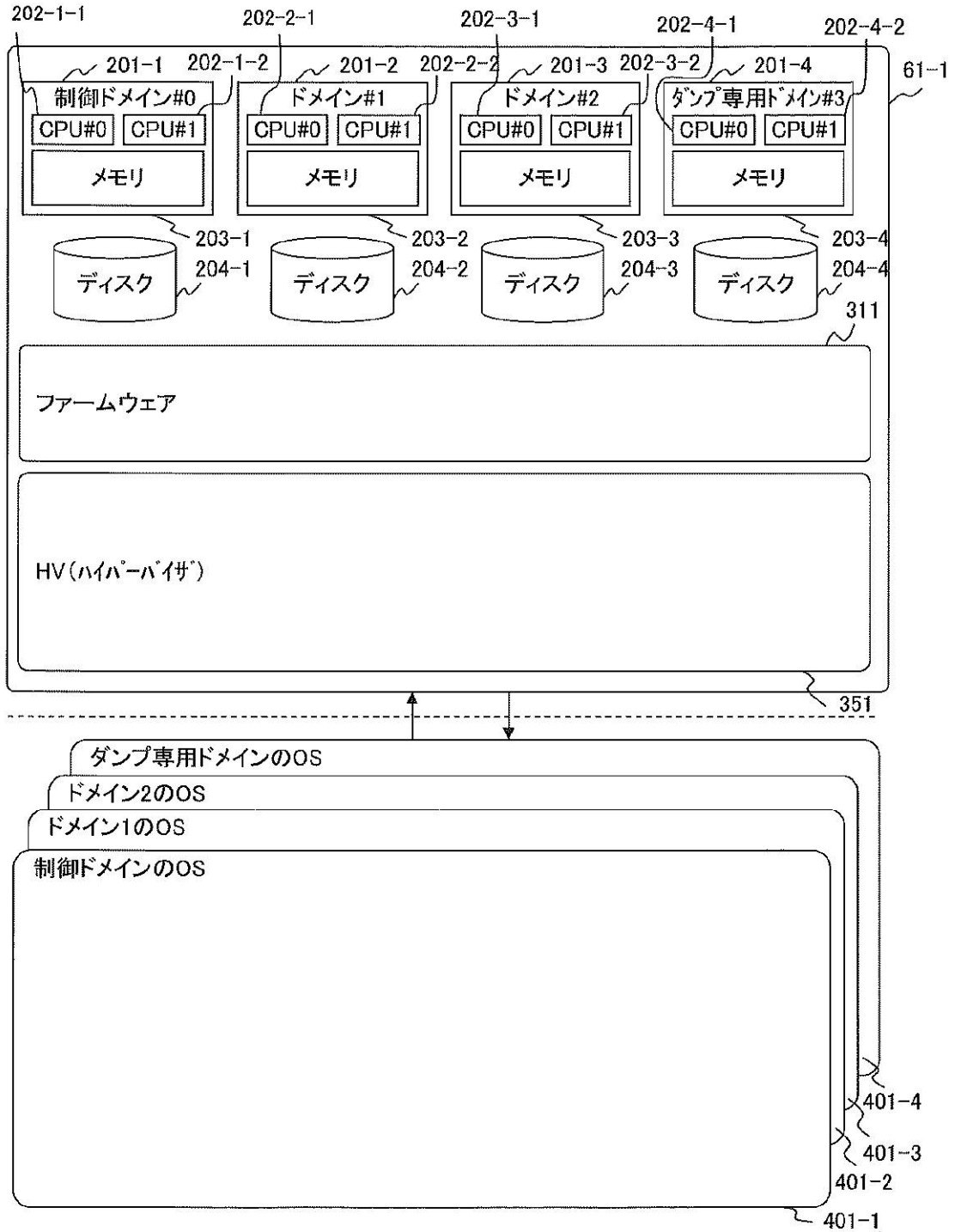
【 図 2 】

実施の形態に係るサーバと機能との対応関係を示す図



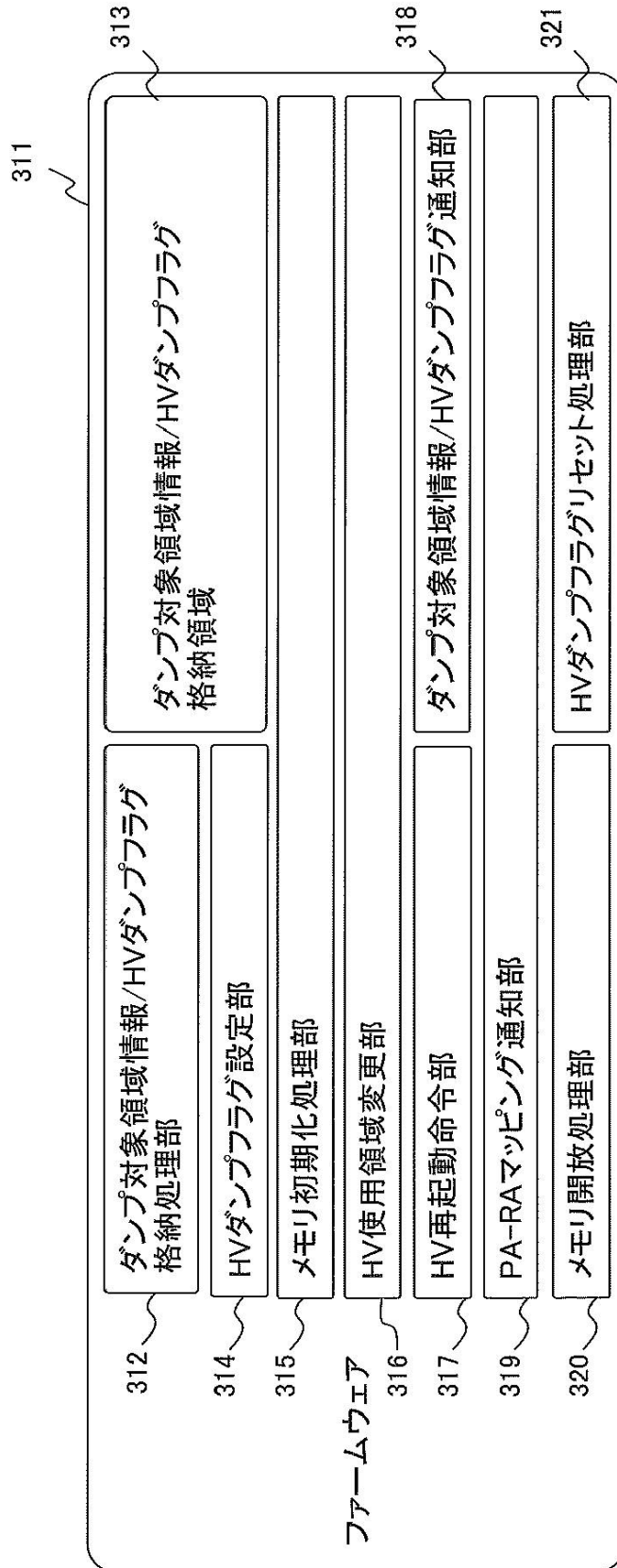
【図3】

実施の形態に係る物理パーティションの機能ブロック図



【 図 4 】

実施の形態に係るファームウェアの構成図



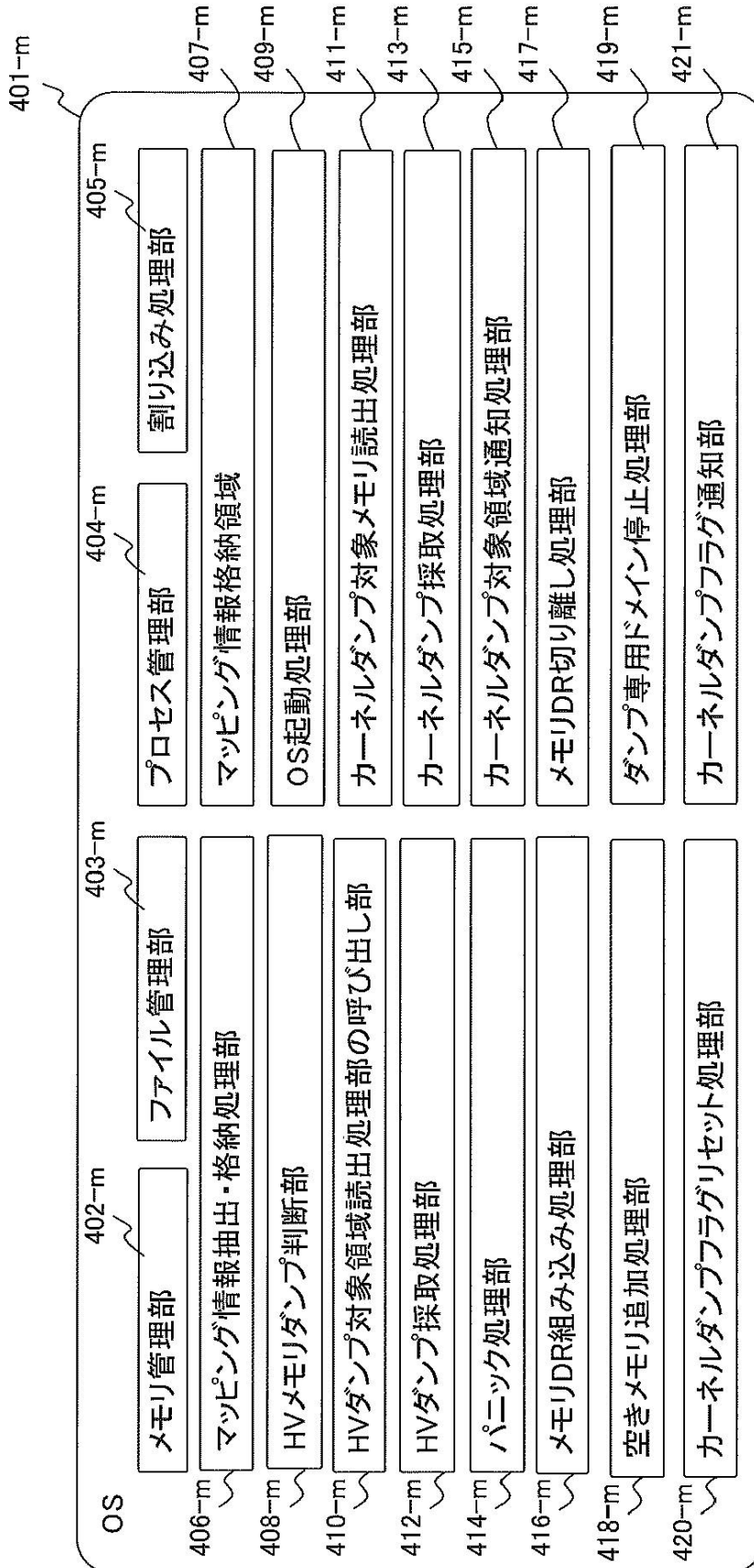
【 図 5 】

実施の形態に係るハイパーバイザの構成図



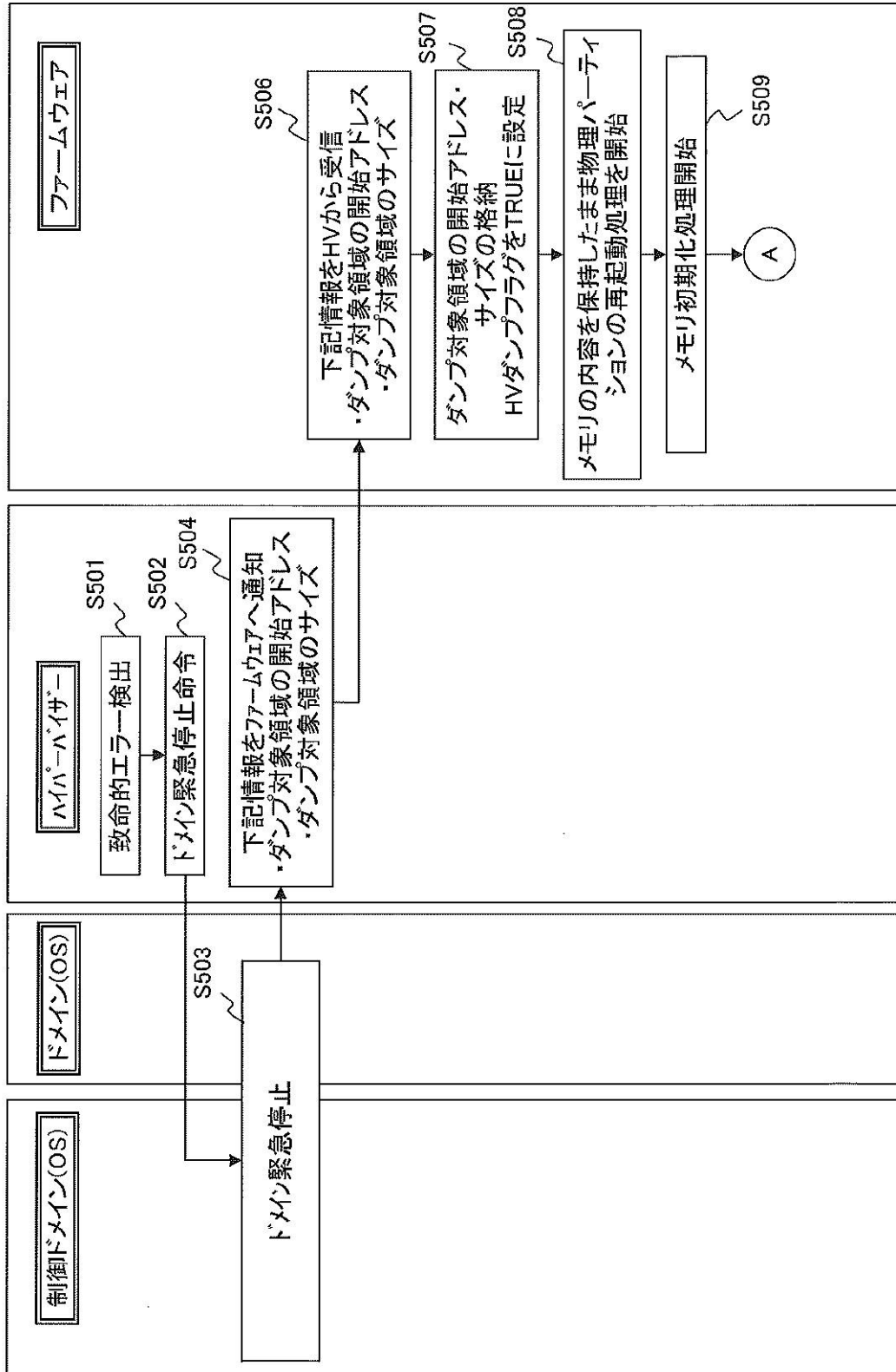
【図6】

実施の形態に係るOSの構成図



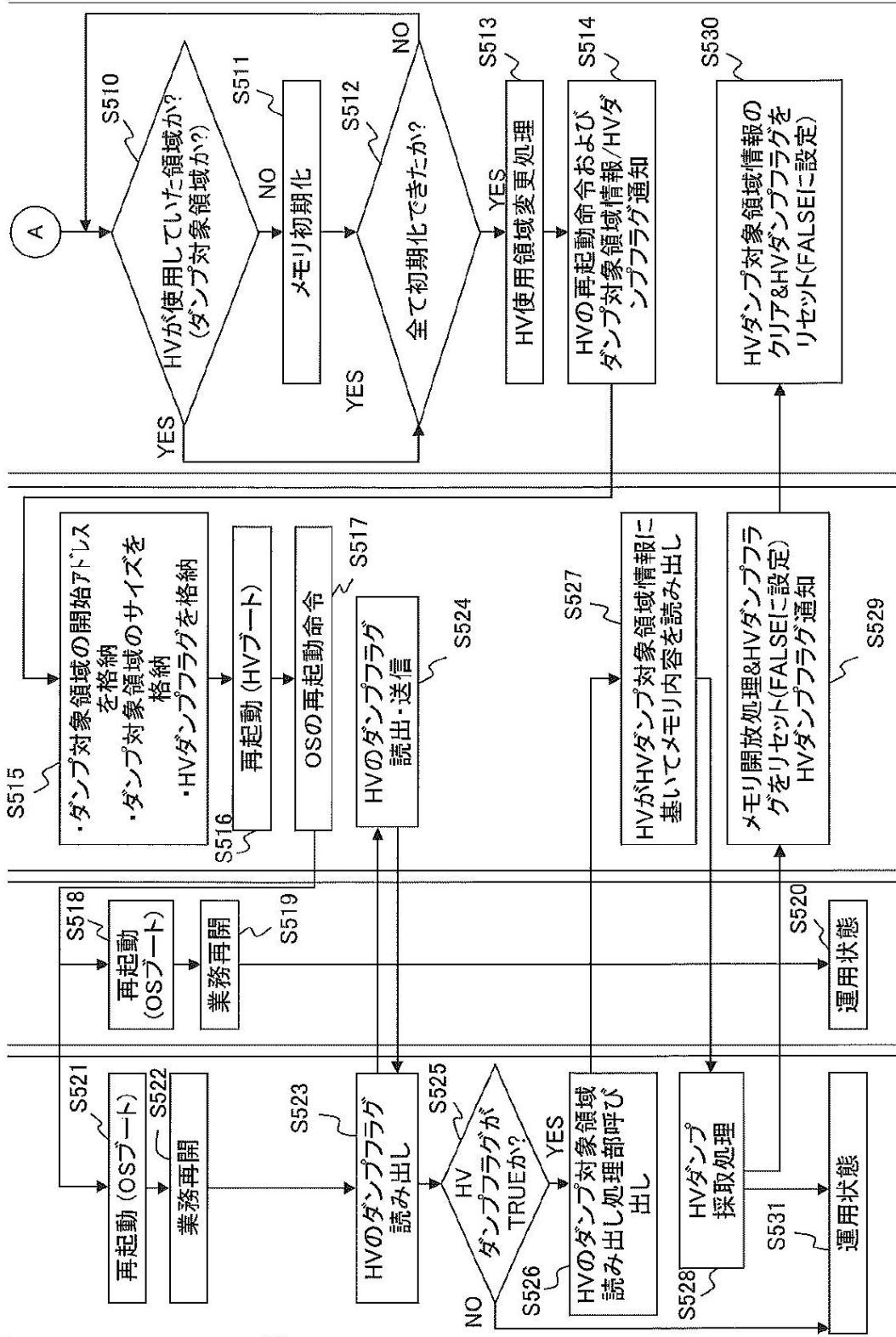
【図7A】

第1の実施の形態に係るメモリダンプ生成処理のフローチャート



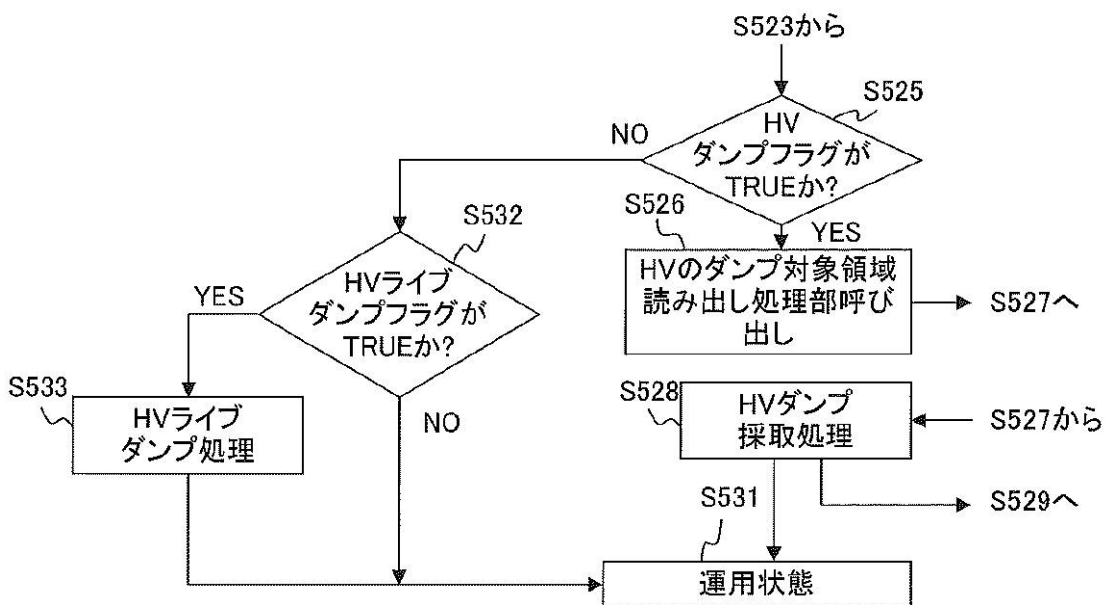
【 図 7 B 】

第1の実施の形態に係るメモリダンプ生成処理のフローチャート



【図7C】

第1の実施の形態に係るメモリダンプ生成処理の変形例のフローチャート



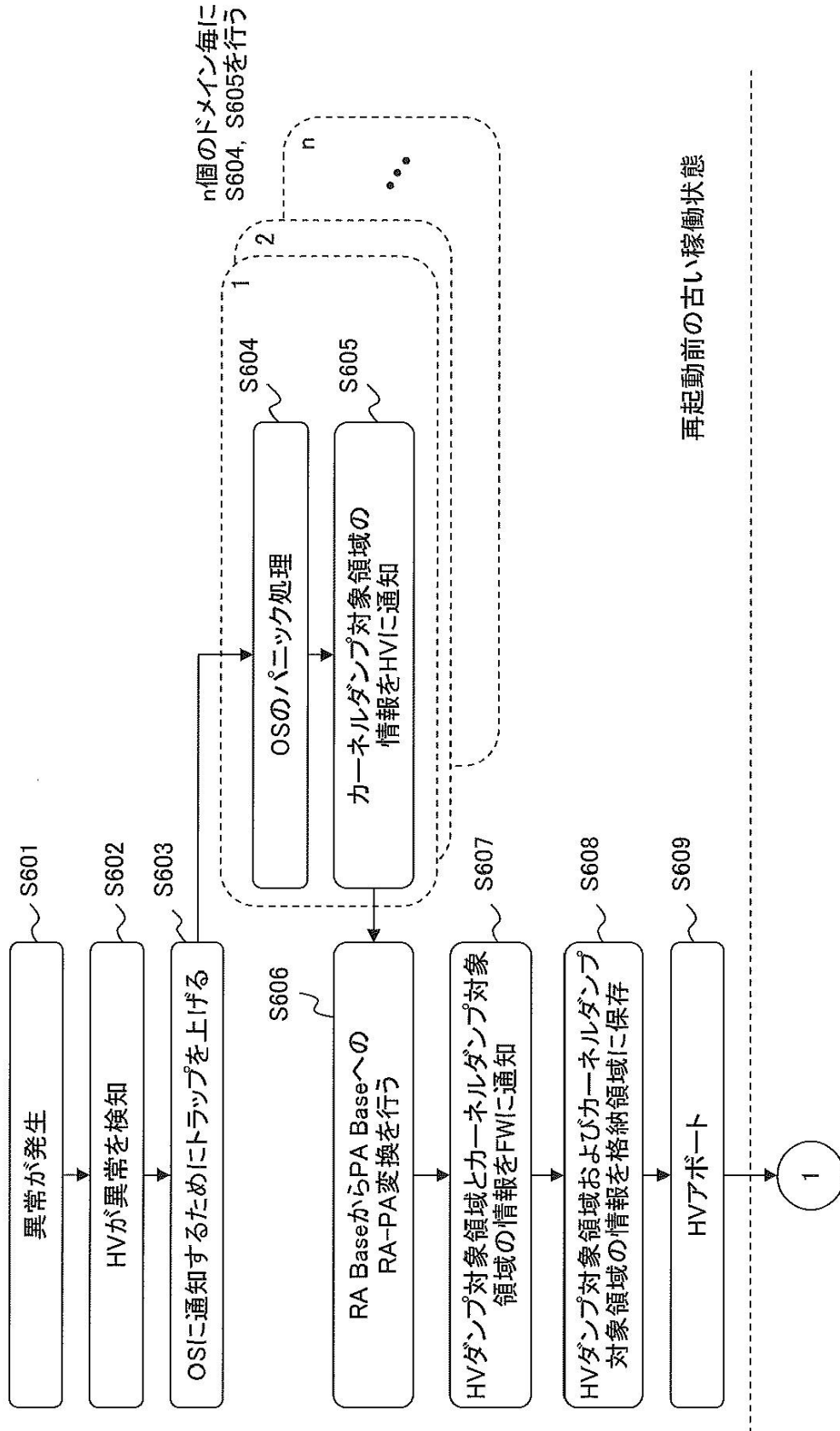
【 図 8 】

HVダンプ対象領域情報の例

Block	PA Base (開始アドレス)	SIZE
0	xxxxxxx	1.5G
1	yyyyyyy	1.0G
N	zzzzzzz	512M

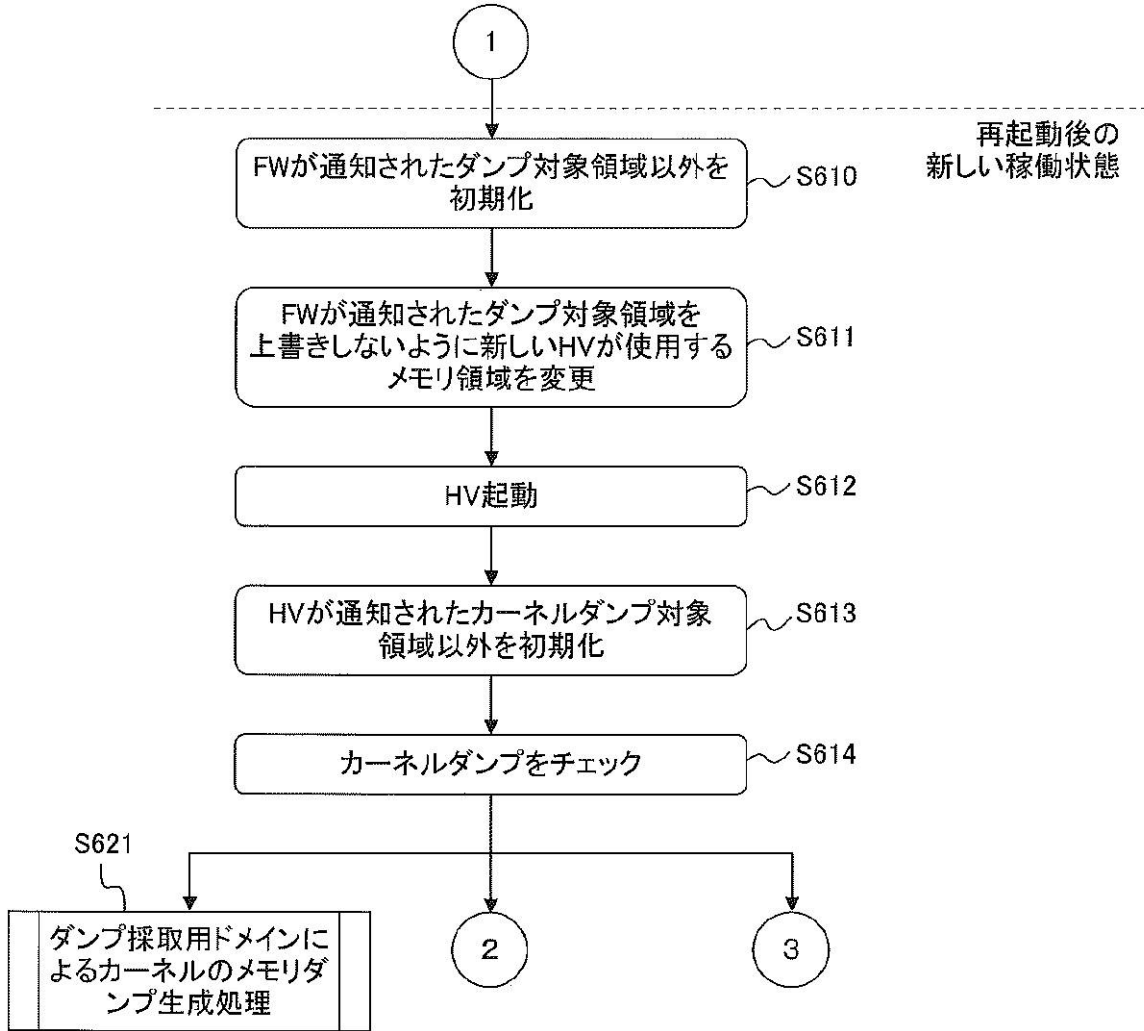
【図9A】

第2の実施の形態に係るメモリダンプ生成処理のフローチャート



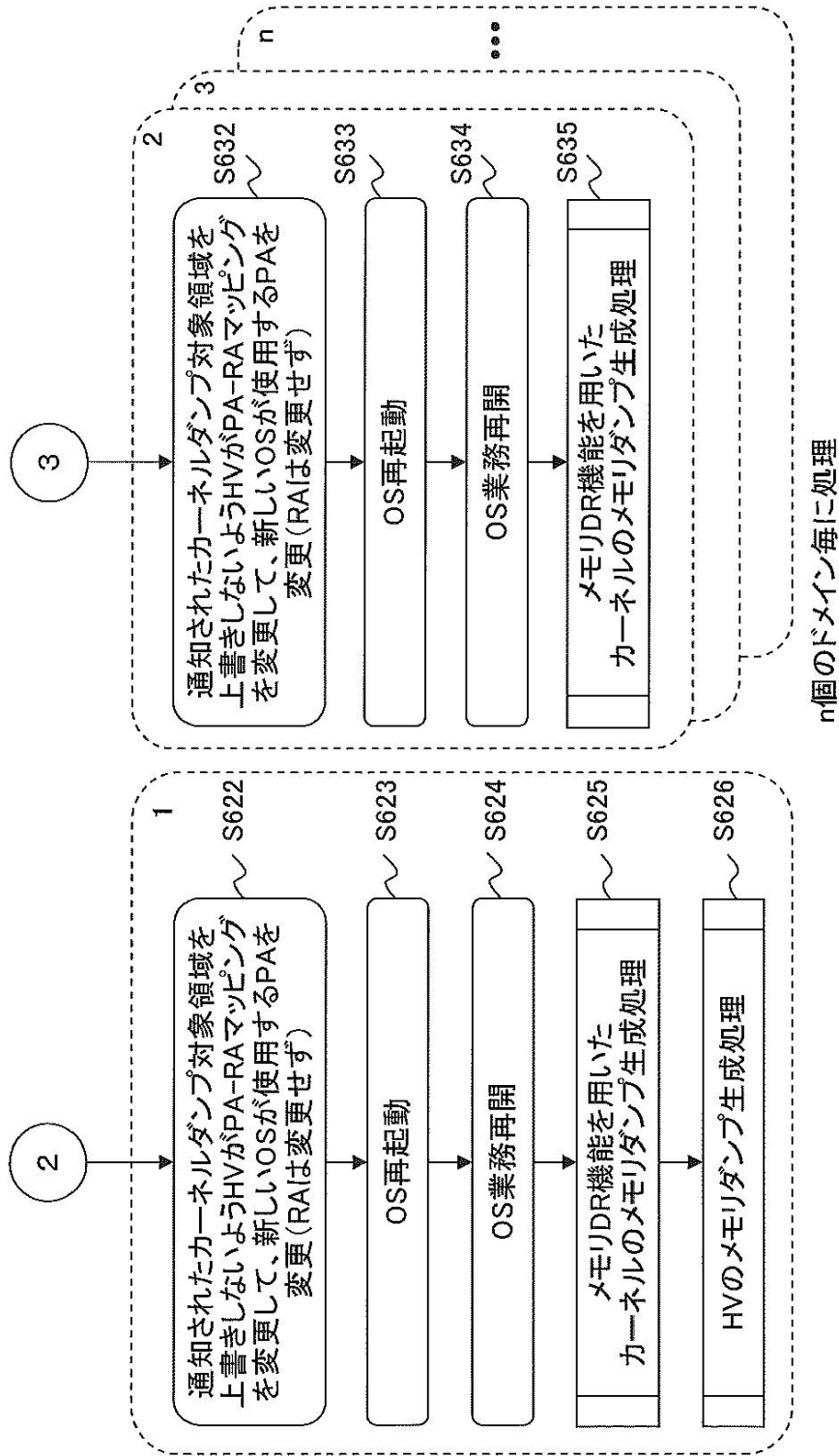
【図9B】

第2の実施の形態に係るメモリダンプ生成処理のフローチャート



【図9C】

第2の実施の形態に係るメモリダンプ生成処理のフローチャート

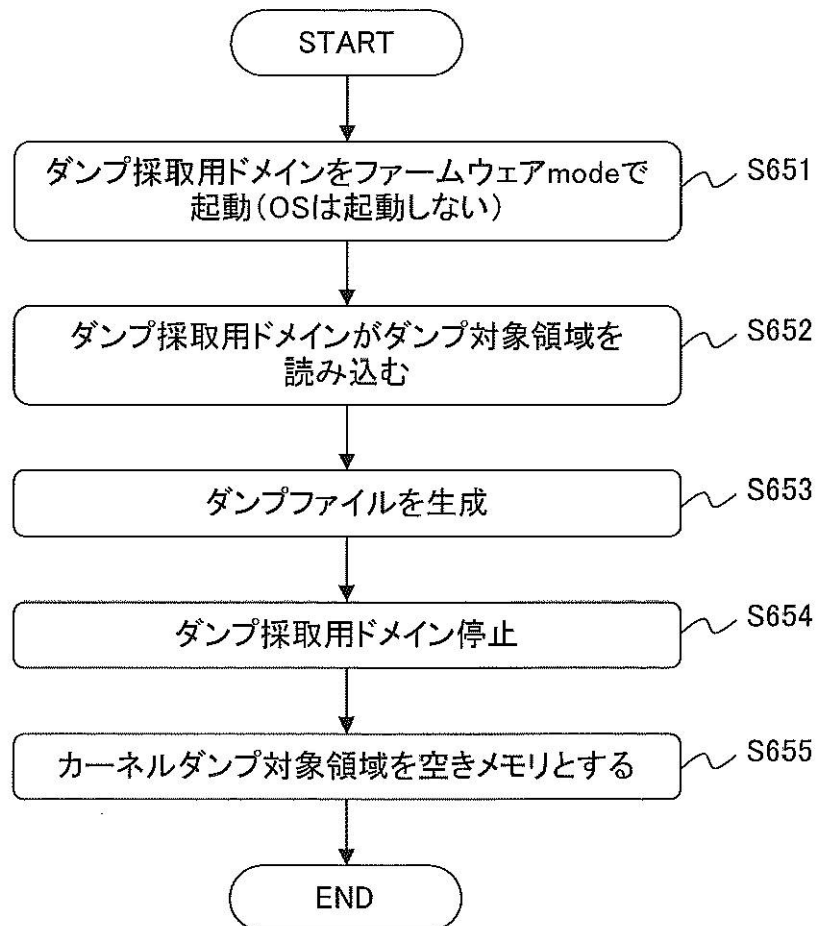


【図 10】

カーネルダンプ対象領域情報の例

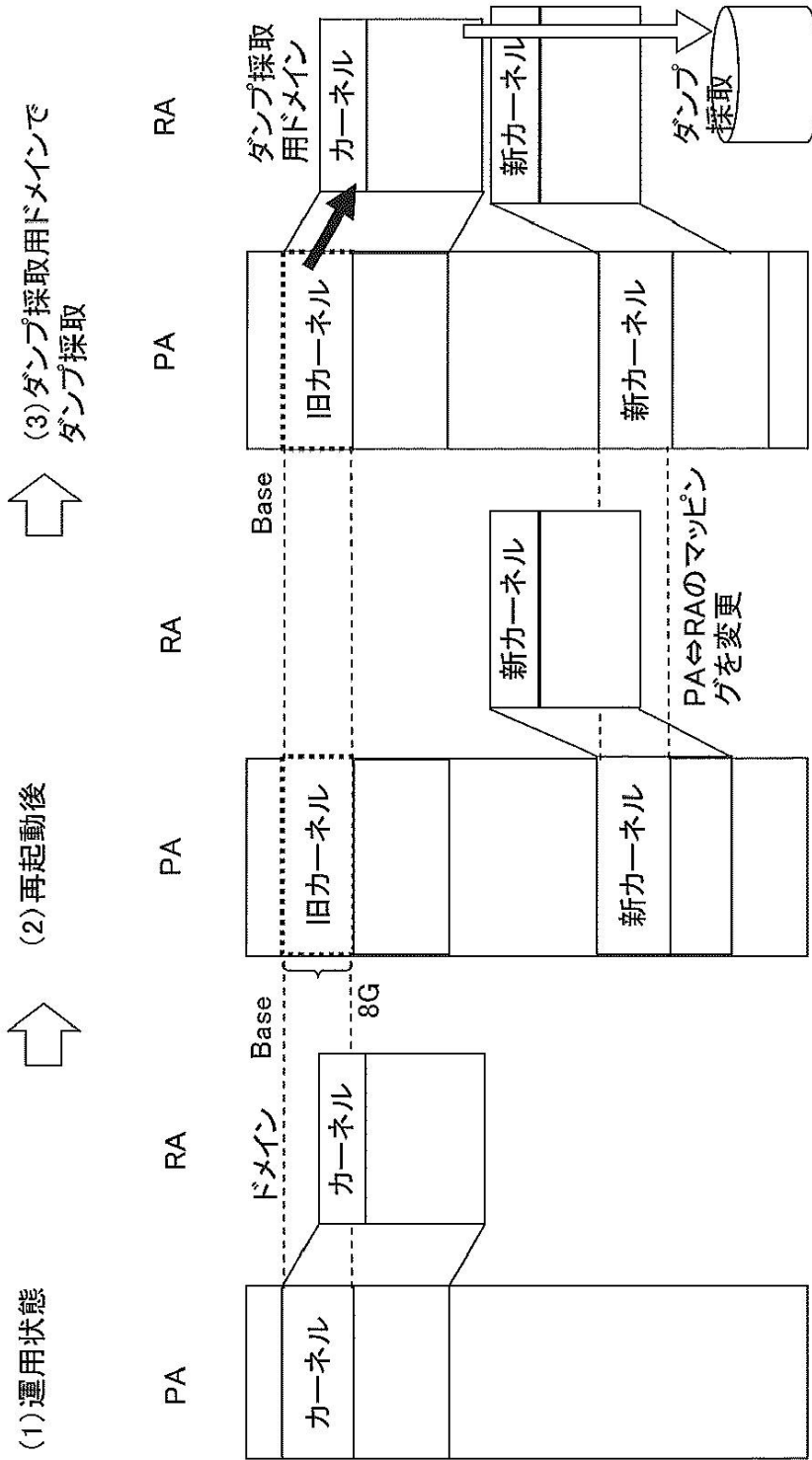
Block	RA Base (開始アドレス)	SIZE
0	aaaaaaaa	1.5G
N	cccccccc	512M

【図 11】

ダンプ採取用ドメインによるカーネルの
メモリダンプ生成処理のフローチャート

【 図 1 2 】

ダンプ採取用ドメインによるメモリダンプの採取を示す図



【 図 1 3 】

ダンプ採取用ドメインによるメモリダンプの採取における PA-RAマッピング情報を示す図

ダンプ後

ドメイン	PA Base (開始アドレス)	SIZE	RA Base (開始アドレス)
#0	yyyyy	8G	aaaaa

ダンプ時

ドメイン	PA Base (開始アドレス)	SIZE	RA Base (開始アドレス)
#3	xxxxx	8G	aaaaa
#0	yyyyy	8G	aaaaa

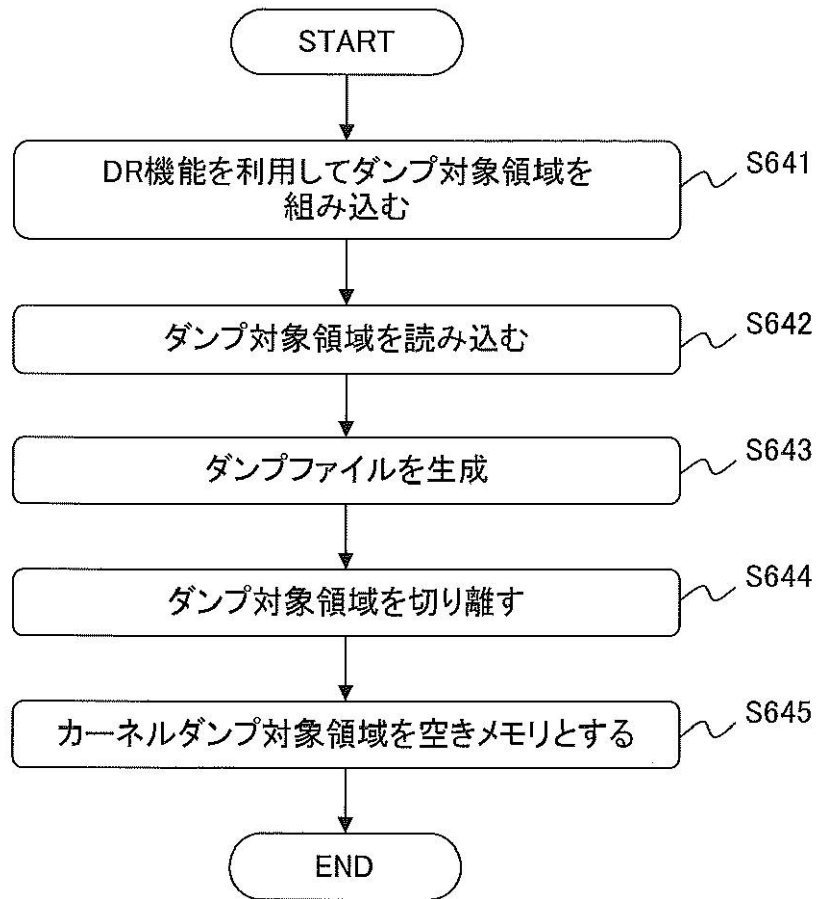
緊急停止時

ドメイン	PA Base (開始アドレス)	SIZE	RA Base (開始アドレス)
#0	xxxxx	8G	aaaaa

ダンプ対象領域

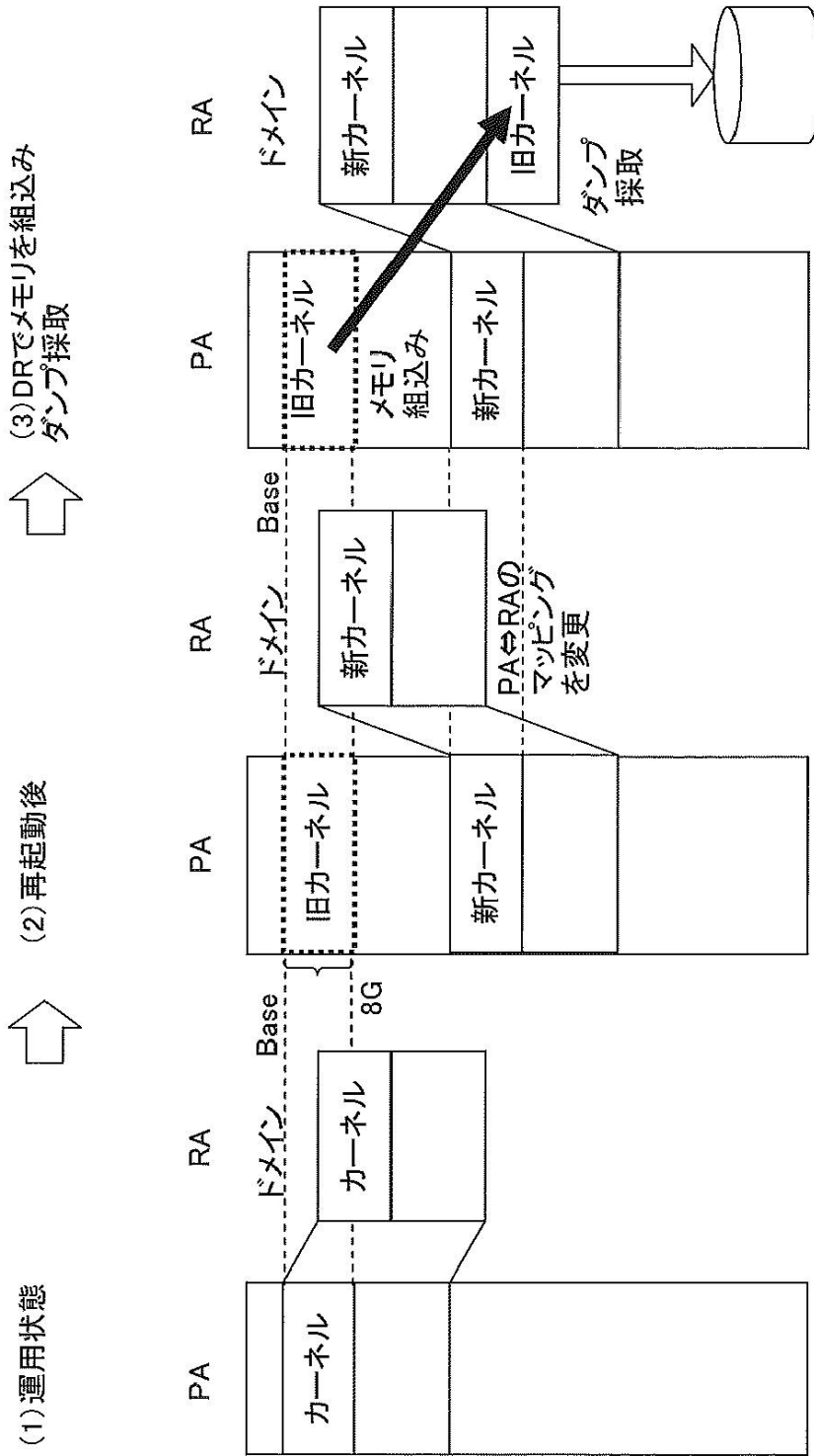
【図14】

メモリDynamic Reconfiguration機能を用いた
カーネルのメモリダンプ生成処理のフローチャート



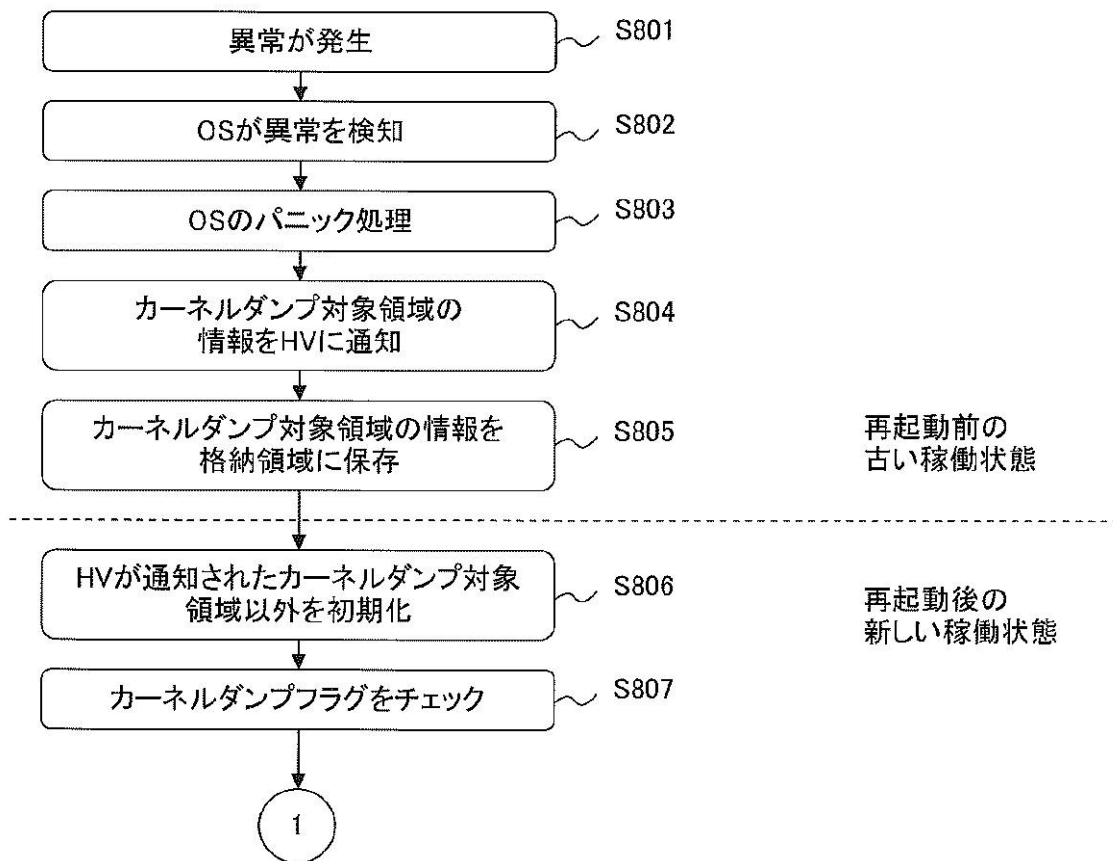
【 図 1 5 】

メモリDynamic Reconfiguration機能を用いた メモリダンプの採取を示す図



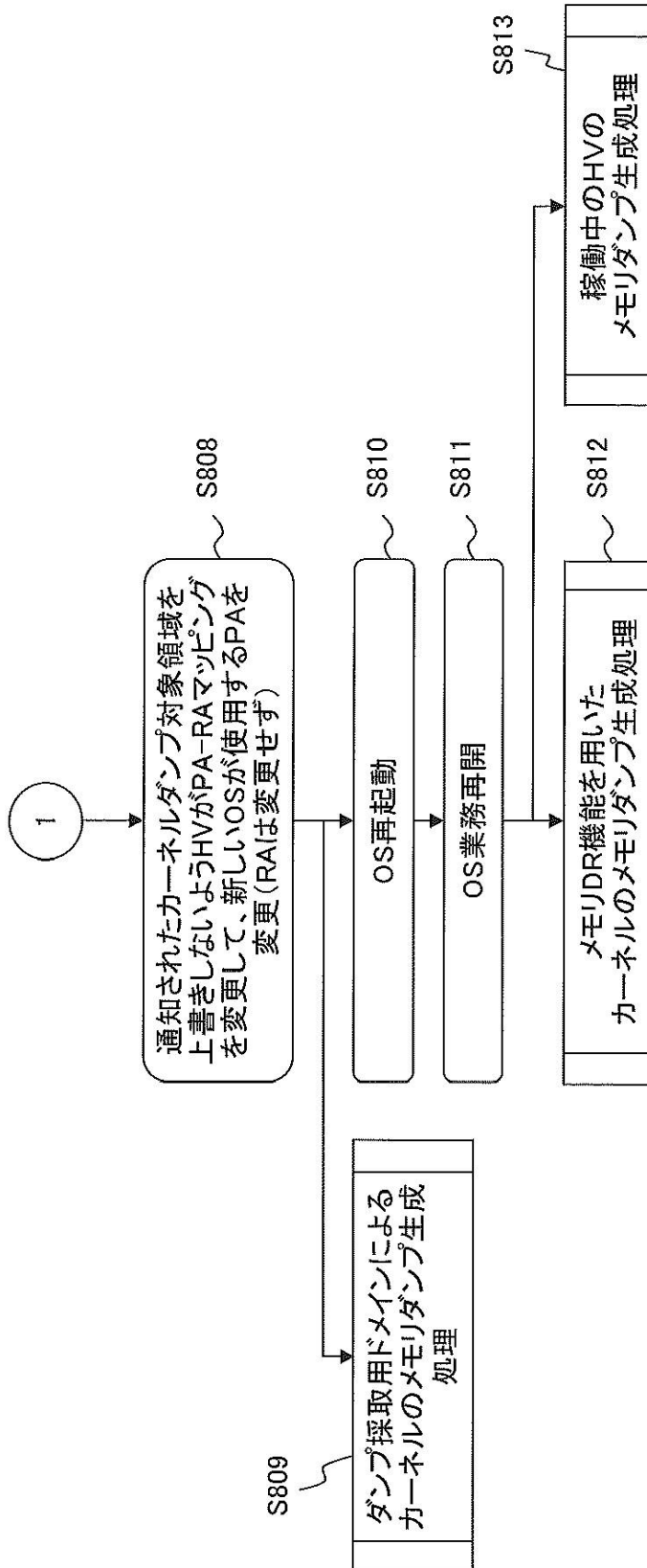
【図18A】

第4の実施の形態に係るメモリダンプ生成処理のフローチャート



【図18B】

第4の実施の形態に係るメモリダンプ生成処理のフローチャート



【図19】

稼働中のハイパーバイザのメモリダンプ生成処理フローチャート

