

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014年2月27日(27.02.2014)



(10) 国際公開番号

WO 2014/030229 A1

- (51) 国際特許分類:
H04L 12/28 (2006.01)
- (21) 国際出願番号: PCT/JP2012/071217
- (22) 国際出願日: 2012年8月22日(22.08.2012)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人(米国を除く全ての指定国について): 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 八田 ゆかり (HATTA, Yukari) [JP/JP]; 〒2591392 神奈川県秦野市堀山下1番地 株式会社日立製作所 I Tプラットフォーム事業本部内 Kanagawa (JP). 早川 典充 (HAYAKAWA, Norimitsu) [JP/JP]; 〒2591392 神奈川県秦野市堀山下1番地 株式会社日立製作所 I Tプラットフォーム事業本部

- 内 Kanagawa (JP). 三木 博史 (MIKI, Hiroshi) [JP/JP]; 〒2591392 神奈川県秦野市堀山下1番地 株式会社日立製作所 I Tプラットフォーム事業本部内 Kanagawa (JP). 野原 史朗 (NOHARA, Shiro) [JP/JP]; 〒2591392 神奈川県秦野市堀山下1番地 株式会社日立製作所 I Tプラットフォーム事業本部内 Kanagawa (JP). 戸塚 崇夫 (TOTSUKA, Takao) [JP/JP]; 〒2591392 神奈川県秦野市堀山下1番地 株式会社日立製作所 I Tプラットフォーム事業本部内 Kanagawa (JP).
- (74) 代理人: 特許業務法人ウィルフォート国際特許事務所 (WILLFORT INTERNATIONAL); 〒1010052 東京都千代田区神田小川町三丁目3番地 神田小川町トーセイビル117階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR,

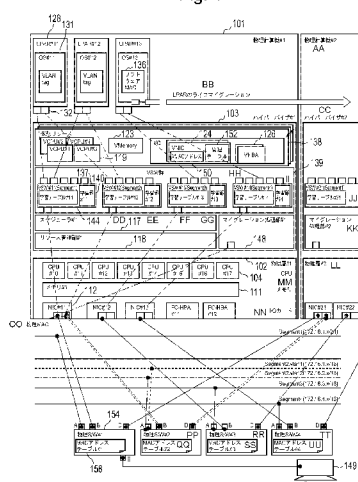
[続葉有]

(54) Title: VIRTUAL COMPUTER SYSTEM FOR RESTORING NETWORK CONNECTION OF VIRTUAL COMPUTER FOR WHICH LIVE MIGRATION HAS BEEN PERFORMED

(54) 発明の名称: ライブマイグレーションされた仮想計算機のネットワーク接続を復元する仮想計算機システム

[図1]

Fig. 1



(57) Abstract: The objective of the present invention is, in a virtual computer system in which communication is performed using a predetermined type of information outside the jurisdiction of a virtualization mechanism, to restore a network connection for a VM (an object VM) for which live migration from a first physical computer to a second physical computer has been performed. In the invention, when a packet is received from a VM, a first virtualization mechanism within a first physical computer extracts a predetermined type of information from the packet, and registers the extracted predetermined type of information in a first management information unit. At the time of live migration, the first virtualization mechanism transmits the predetermined type of information within the first management information unit to a second virtualization mechanism within a second physical computer. The second virtualization mechanism registers the predetermined type of information in a second management information unit, and in order to restore a network connection of the object VM, transmits a predetermined type of packet that contains the predetermined type of information from at least one among one or more physical I/O devices of the second physical computer.

(57) 要約:

[続葉有]

- | | |
|----------------------------------|---------------------------------|
| 101 Physical computer #1 | DD Learning table #12 |
| 102 Physical layer #1 | EE Learning unit #12 |
| 103 Hypervisor #1 | FF Learning table #13 |
| 111 Memory #1 | GG Learning unit #13 |
| 117 Scheduler #1 | HH Learning table #14 |
| 118 Resource management unit #1 | II Learning unit #14 |
| 136 Software MAC | JJ Learning table #21 |
| 138 Virtual resources | KK Migration processing unit #2 |
| 139 VSW group | LL Physical layer #2 |
| 140 Learning unit #11 | MM Memory |
| 144 Learning table #11 | NN IO card |
| 148 Migration processing unit #1 | OO Physical MAC |
| 149 Management console | PP Physical SW #2 |
| 150 VMAC address | QQ MAC address table #2 |
| 152 Management table | RR Physical SW #3 |
| 154 Physical SW #1 | SS MAC address table #3 |
| 158 MAC address table #1 | TT Physical SW #4 |
| AA Physical computer #2 | UU MAC address table #4 |
| BB Live migration of LPAR | |
| CC Hypervisor #2 | |



LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

パ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

仮想化機構の管轄外の所定種類の情報を使用した通信が行われる仮想計算機システムにおいて第 1 の物理計算機から第 2 の物理計算機にライブマイグレーションされた対象の VM (対象 VM) のネットワーク接続を復元する。第 1 の物理計算機内の第 1 の仮想化機構が、VM からパケットを受信した場合、パケットから、所定種類の情報を抽出し、抽出した所定種類の情報を第 1 の管理情報ユニットに登録するようになっており、ライブマイグレーションのときに、第 1 の管理情報ユニット中の所定種類の情報を、第 2 の物理計算機内の第 2 の仮想化機構に送信する。第 2 の仮想化機構は、その所定種類の情報を第 2 の管理情報ユニットに登録し、対象 VM のネットワーク接続を復元するべく、所定種類の情報を含んだ所定種類のパケットを、第 2 の物理計算機の 1 以上の物理 I/O デバイスの少なくとも 1 つから送信する。

明 細 書

発明の名称：

ライブマイグレーションされた仮想計算機のネットワーク接続を復元する
仮想計算機システム

技術分野

[0001] 本発明は、仮想計算機のライブマイグレーションに関する。

背景技術

[0002] 物理計算機が有するハイパーバイザのような仮想化機構により、物理計算機に 1 又は複数の仮想計算機が構築される。稼働中の仮想計算機を第 1 の物理計算機から第 2 の物理計算機にマイグレーションすることであるライブマイグレーション（以下、「LM」と言うことがある）の技術が知られている（例えば特許文献 1）。

[0003] 以下、LM元の物理計算機を「第 1 の物理計算機」と言い、第 1 の物理計算機が有する仮想化機構を「第 1 の仮想化機構」と言い、LM先の物理計算機を「第 2 の物理計算機」と言い、仮想的なNIC（Network Interface Card）を「VNIC」と言い、第 2 の物理計算機が有する仮想化機構を「第 2 の仮想化機構」と言い、LM対象の仮想計算機（仮想マシン）を「対象VM」と言い、第 1 の仮想化機構が管理するVNICであってLM前に対象VMに割り当てられていたVNICを「第 1 のVNIC」と言い、第 2 の仮想化機構が管理するVNICであってLM後に対象VMに割り当てられるVNICを「第 2 のVNIC」と言うことがある。

先行技術文献

特許文献

[0004] 特許文献1：特開2012-108816号公報

発明の概要

発明が解決しようとする課題

- [0005] 対象VMの上位マシン（例えば、対象VMにアクセスする物理的又は仮想的な計算機、或いは、対象VM上で動作する仮想的な計算機）がいたとしても、LMでは、上位マシンに気づかれることなく、対象VMをマイグレーションする必要がある。このため、LM後に対象VMと外部環境との接続を復元する必要がある。
- [0006] 外部環境のネットワーク接続情報として、例えば、PSW（物理的なネットワークスイッチ）が保持しているMACアドレステーブルがある。MACアドレステーブルは、宛先となるMACアドレスが割り当てられているマシンがPSWのどのポートに接続されているかを表す情報を有する。
- [0007] LMの後、MACアドレステーブルを更新する必要がある。この更新は、第2の仮想化機構が、RARP（Reverse Address Resolution Protocol）パケットを発行することで行われる。RARPは、MACアドレスを用いてIPアドレスを取得するためのプロトコルである。
- [0008] 具体的には、第1及び第2の仮想化機構の間で、第1のVNICのMACアドレス#Xと第2のVNICのMACアドレス#Yが交換される。第2の仮想化機構は、第2の仮想化機構が有することになったMACアドレス#Xを指定したRARPパケットをブロードキャストする。ネットワーク上のRARPサーバが、そのRARPパケットを受信し、そのRARPパケットに応答して、MACアドレス#Xに対応するIPアドレス#Kを第2の仮想化機構に返す。これにより、第2の物理計算機がIPアドレス#Kを使用した通信をすることができるようになる。RARPパケットは、PSWで中継されるが、PSWは、MACアドレステーブルにおけるMACアドレス#Xに対応したポートの識別情報を、そのRARPパケットを受信したポートの識別情報に更新する。これにより、MACアドレス#Xとポートとの対応関係が更新され、LM後の対象VMについて、ネットワークの接続が復元することになる。
- [0009] しかし、LM前において、第1の仮想化機構の管轄の外で、第1のVNICのMACアドレスとは異なるMACアドレスが、対象VMによる通信にお

いて使用されることがある。

[0010] その一例として、Teamingと呼ばれる技術が適用されている環境がある。Teamingとは、計算機に割り当てられた1以上のNICの複数のポートを束ねて1つのポートとして運用し負荷分散、帯域向上及び耐障害性の向上を図るための技術である。この技術がVNICに適用された場合、1以上のVNICに係る複数のポート（チーム設定された複数のポート）は一体として運用され、下記のうちの少なくとも1つが行われる。

（＊）複数のポートのうちの第1のポートに障害が発生した場合に、第1のポートが論理的に切り離され、複数のポートのうちの第2のポートを使用して通信が継続される。

（＊）複数のポートを介して複数の通信相手と通信する際に分担して送受信が行われることで、帯域幅が増加され、又は、負荷が分散される。

[0011] Teamingで使用されるMACアドレスは、VNICに割り当てられているMACアドレスと異なるMACアドレスであることがある。仮想計算機のOS（ゲストOS）が使用するMACアドレスの変更は、仮想化機構とは独立して実施される。

[0012] そして、PSWが有するMACアドレステーブルには、ゲストOSが使用するMACアドレスが登録される。具体的には、例えば、PSWは、外部から来たパケットを受信した場合、受信したパケットに含まれているMACアドレスがMACアドレステーブルに無ければ、そのMACアドレスと、そのパケットを受信したポート（PSWのポート）のIDとをMACアドレステーブルに登録するようになっている。

[0013] そのため、LM後に第2の仮想化機構が送信するRARPパケットにおいて、正しいMACアドレス（具体的には、PSWのMACアドレステーブルに登録されているMACアドレスと同じMACアドレス）を指定することができないケースがあり得る。この場合、対象VMのネットワーク接続を復元することはできない。

[0014] 上記のことは、MACアドレスに限らず、第1の仮想化機構が管理する情

報とは異なる情報が、対象VMによる通信において使用される他のケースについてもあり得る。例えば、対象VMが `tag VLAN` (Virtual LAN) を使用するケースである (LANは、Local Area Networkの略称)。`tag VLAN`を経由する通信では、`VLAN tag` (`VLAN ID`等を含んだ情報) が使用されるが、仮想化機構は、対象VMが使用する `VLAN tag` を知らない。このようなケースでは、ゲストOSのMACアドレスがVNICのMACアドレスと同じであったとしても、上記のように、LM後に対象VMのネットワーク接続を復元することができないという問題が生じ得る。

[0015] このようなことを回避する1つの方法として、対象VMが使用するMACアドレス或いは `VLAN tag` のような情報を仮想化機構に管理者が管理コンソールを通じて手動で入力する方法が考えられる。しかし、誤った情報が入力される或いは情報の入力漏れが生じるおそれがある。この場合、LM後に対象VMのネットワーク接続を復元することはできない。

[0016] また、仮想計算機或いはその上位マシンがどのように運用されているかを仮想化機構の管理者が管理することは困難であるケースがある。そのようなケースの一例として、クラウド環境で `PaaS` (Platform as a Service) を提供するケースである。通常、仮想化機構の管理者は、仮想計算機或いはその上位マシンがどのように運用されているかを把握していない。このため、仮想計算機が使用する情報を管理者が手動で入力する場合、仮想マシン或いはその上位マシンの使用者と仮想化機構の管理者の連携が必要になる。

課題を解決するための手段

[0017] 第1の仮想化機構は、第1の管理情報ユニットを有し、第2の仮想化機構が、第2の管理情報ユニットを有する。第1の仮想化機構が、仮想計算機からパケットを受信した場合、パケットから、所定種類の情報を抽出し、抽出した所定種類の情報を第1の管理情報ユニットに登録する。第1の仮想化機構は、LMのときに、第1の管理情報ユニット中の所定種類の情報を、第2の仮想化機構に送信する。第2の仮想化機構は、その所定種類の情報を第2

の管理情報ユニットに登録し、LMされた仮想計算機のネットワーク接続を復元するべく、所定種類の情報を含んだ所定種類のパケットを、第2の物理計算機の1以上の物理I/Oデバイスの少なくとも1つから送信する。

発明の効果

[0018] 仮想化機構の管轄外の情報を使用した通信が行われる環境であっても、仮想化機構の管理者によって当該管轄外情報が手動で入力されること無しに、ライブマグレーションされた仮想計算機のネットワーク接続を復元することができる。仮想化機構の管理者が手動で管轄外情報を入力する必要が無いので、当該情報入力のために仮想計算機或いはその上位マシンの使用者と仮想化機構の管理者が連携する必要も無い。

図面の簡単な説明

[0019] [図1]図1は、実施形態に係る仮想計算機システムとその外部環境の一例を示す。

[図2]図2は、VNIC#11の管理テーブル#11とVNIC#12の管理テーブル#12の例を示す。

[図3]図3は、tag VLANを使用していないLPAR128から送信されるパケットの構成例を示す。

[図4]図4は、tag VLANを使用しているLPAR128から送信されるパケットの構成例を示す。

[図5]図5は、LM対象のLPAR128がtag VLANを使用していないLPAR128である場合に送信されるRARPパケットの構成例を示す

[図6]図6は、ハイパーバイザ103内のVSW137が行う処理の流れを示す。

[図7]図7は、MACアドレス学習テーブル144の構成例を示す。

[図8]図8は、メモリ111のうちハイパーバイザ103に関する情報が占める領域を示す。

[図9]図9は、図6のS603（MACアドレス学習処理）の流れを示す。

[図10]図10は、学習データ収集処理の流れを示す。

[図11]図 1 1 は、LM処理の流れを示す。

[図12]図 1 2 は、LM元のハイパーバイザ# 1 におけるVSW# 1 1 のLM前後のMACアドレス学習テーブル# 1 1 の一例を示す。

[図13]図 1 3 は、LM先のハイパーバイザ# 2 におけるVSW# 2 1 のLM前後のMACアドレス学習テーブル# 2 1 の一例を示す。

[図14]図 1 4 は、PSW# 1 におけるLM前のMACアドレステーブル# 1 の一例を示す。

[図15]図 1 5 は、PSW# 1 におけるLM前のMACアドレステーブル# 1 の一例を示す。

発明を実施するための形態

[0020] 以下、図面を参照して、一実施形態を説明する。

[0021] その際、同種の複数の要素を区別しないで説明する際には、その要素の参照符号を使用し、同種の要素を区別して説明する際には、その要素の参照符号に代えてその要素に付した通し番号を使用する。例えば、複数の物理計算機を区別しないで説明する際には、「物理計算機 1 0 1」と記載し、複数の物理計算機を区別して説明する際には、「物理計算機 # 1」、「物理計算機 # 2」と記載する。

[0022] また、以下の説明では、a a a情報の一例として、a a aテーブルという名称で説明するが、情報はテーブルの形式に限らず他種の形式でも良い。

[0023] 図 1 は、実施形態に係る仮想計算機システムとその外部環境の一例を示す。

[0024] 仮想計算機システムは、複数の物理計算機 1 0 1 を含む。仮想計算機システムの外部環境は、例えば、複数のPSW 1 5 4 と、管理コンソール 1 4 9 とを含む。

[0025] 物理計算機 1 0 1 は、物理層（物理リソース） 1 0 2 を有する。物理層 1 0 2 は、1 以上のCPU 1 0 4、1 以上のメモリ 1 1 1、及び1 以上の物理 I / O（Input/Output）デバイスを含む。物理 I / O デバイスは、例えば、通信ネットワークを介した通信のためのインタフェースデバイスであり、物

理 I/O デバイスが通信ネットワークに接続されていてよい。1 以上の物理 I/O デバイスは、例えば、3 つの物理的な NIC 112 と、2 つの FC (Fibre Channel) - HBA (Host Bus Adaptor) 115 である。図示の例では、物理計算機 # 1 に関して、NIC # 11 の第 1 のポートに、PSW # 1 のポート # A が接続されており、NIC # 12 の第 1 のポートに、PSW # 3 のポート # B が接続されており、NIC # 13 の第 1 のポートに、PSW # 4 のポート # B が接続されている。物理計算機 # 2 に関して、NIC # 21 の第 1 のポートに PSW # 1 のポート # D が接続されており、NIC # 22 の第 1 のポートに PSW # 3 のポート # D が接続されている。

[0026] 物理計算機 101 において、物理層 102 の少なくとも 1 つの CPU 104 においてコンピュータプログラム（例えばファームウェア）が実行されることで、ハイパバイザ 103 が実現される。ハイパーバイザ 103 は、物理層 102 (CPU 104、メモリ 111、I/O デバイス (NIC 112 及び FC-HBA 115)) を論理的に分割し、且つ、論理的に分割することにより得られた複数の仮想リソースをそれぞれ複数の LPAR (Logical Partition) に割り当てることにより、1 台の物理計算機 101 上に複数の仮想計算機を動作させることを可能とする機構である。仮想リソースの数は LPAR の数と同じであり、仮想リソースは LPAR に 1 対 1 で割り当てられる。仮想リソースが LPAR に割り当てられることで仮想計算機が実現されるので、以下、本実施形態では、仮想計算機を「LPAR」と呼ぶ。1 つの LPAR 128 に割り当てられる 1 つの仮想リソースは、例えば、1 以上の CPU 104 に基づく 1 以上の VCPU (仮想 CPU) 119、1 以上のメモリ 111 に基づく 1 以上の VMEM (仮想メモリ) 123、及び、1 以上の I/O デバイスに基づく 1 以上の仮想 I/O デバイスを含む。1 以上の仮想 I/O デバイスは、例えば、1 以上の NIC 112 に基づく 1 以上の VNIC 124、及び/又は、1 以上の FC-HBA 115 に基づく 1 以上の VHBA (仮想 HBA) 126 である。

[0027] LPAR 128 では、OS (ゲスト OS) 131 が実行される。ゲスト OS

S 1 3 1 は、パケットを送信する際、パケットに、所定種類の情報を含めるようになっている。所定種類の情報は、例えば、ゲスト O I S 1 3 1 が使用する M A C アドレス、及び／又は、V L A N t a g を含む。V L A N t a g は、ゲスト O S 1 3 1 の通信において t a g V L A N が使用される場合に、ゲスト O S 1 3 1 が送信するパケットに含まれる。ゲスト O S 1 3 1 が送信するパケットは、そのゲスト O S 1 3 1 を実行する L P A R 1 2 8 に割り当てられている V N I C 1 2 4 又は V H B A 1 2 6 から送信される。そして、パケットに含まれる M A C アドレスは、ゲスト O S 1 3 1 が独自に使用する M A C アドレス（以下、「ソフトウェア M A C」と言うことがある）であって、V N I C の M A C アドレスとは異なる M A C アドレスであることがある。

[0028] L P A R 1 2 8 から送信されたパケットは、複数の P S W 1 5 4 のうちの少なくとも 1 つが受信する。P S W 1 5 4 は、物理的なネットワークスイッチである。P S W 1 5 4 は、複数のポート（物理ポート）と、M A C アドレステーブル 1 5 8（正確には、M A C アドレステーブル 1 5 8 を記憶するメモリのような記憶デバイス）と、図示しないプロセッサとを有している。M A C アドレステーブル 1 5 8 は、そのテーブル 1 5 8 を有する P S W 1 5 4 の物理ポートの I D と M A C アドレスとの対応関係を表すアドレス管理情報の一例である。ここで言う「対応関係」は、どの物理ポートにどのゲスト O S が接続されているかを表す。そのため、M A C アドレステーブル 1 5 8 に登録される物理ポート I D は、パケットの転送先物理ポートの I D と言うこともできる。M A C アドレステーブル 1 5 8 は、例えば、図 1 4 に示すように、M A C アドレスと、V L A N t a g の少なくとも一部分の一例である V L A N I D と、物理ポートの I D との組を有する（図 1 4 は、P S W # 1 の M A C アドレステーブル # 1 の一例である。

[0029] P S W # 1（P S W # 1 のプロセッサ）は、パケットを受信した場合、対象 M A C アドレス（受信したパケットに含まれている M A C アドレス）が M A C アドレステーブル # 1 に無ければ、対象 M A C アドレスと対象 V L A N

ＩＤ（パケットに含まれているＶＬＡＮ ｔａｇ中のＶＬＡＮ ｌｄ）と、対象物理ポート（パケットを受信した物理ポート）のＩＤとの組をＭＡＣアドレステーブル＃１に登録するようになっている。対象ＶＬＡＮ ｌｄが無い場合には、ＭＡＣアドレステーブル＃１には、ＶＬＡＮ ｌｄとして、指定無しを意味する値が登録される。また、ＰＳＷ＃１（ＰＳＷ＃１のプロセッサ）は、対象ＭＡＣアドレスがＭＡＣアドレステーブル＃１に有るが、対象ＭＡＣアドレスに対応している物理ポートＩＤが対象物理ポートＩＤと異なっていれば、ＭＡＣアドレステーブル＃１における、対象ＭＡＣアドレスに対応した物理ポートＩＤを、対象物理ポートＩＤに更新するようになっている。また、ＰＳＷ＃１（ＰＳＷ＃１のプロセッサ）は、パケットを受信した場合、対象ＭＡＣアドレスがＭＡＣアドレステーブル＃１に有り、且つ、対象ＭＡＣアドレスに対応しているＶＬＡＮ ｌｄが、対象ＶＬＡＮ ｌｄと異なっていれば、新たなエントリを追加し、そのエントリに、対象ＭＡＣアドレスと対象ＶＬＡＮ ｌｄを登録する。ＰＳＷ＃１のこのような機能により、ＭＡＣアドレステーブル＃１には、ゲストＯＳが使用するＭＡＣアドレス（及びＶＬＡＮ ｔａｇ）が登録されるようになっており、故に、そのＭＡＣアドレスが、ソフトウェアＭＡＣであれば、ＬＰＡＲ１２８のＶＮＩＣ１２４のＭＡＣアドレスではなく、ソフトウェアＭＡＣが登録される。なお、ＰＳＷ＃１のこのような機能と同じ機能を、他のＰＳＷ１２８も有している。

[0030] ハイパーバイザ１０３は、スケジューラ１１７、リソース管理部１１８、マイグレーション処理部１４８、複数のＶＳＷ（仮想スイッチ）１３７、及び、前述した１以上の仮想リソースを有する。

[0031] スケジューラ１１７は、物理層１０２の論理分割を行い、リソース管理部１１８は、物理層１０２としての物理リソース等のリソースを管理する。スケジューラ１１７及びリソース管理部１１８により、物理層１０２（物理リソース）を論理的に分割することにより得られた複数の仮想リソース（例えば、ＶＣＰＵ１１９、ＶＭＥＭ１２３、ＶＮＩＣ１２４及びＶＨＢＡ１２６

）をそれぞれ複数のL P A R 1 2 8に割り当てることができる。

[0032] V N I C 1 2 4 は、ハイパーバイザ 1 0 3 が割り当てた仮想MACアドレスを持つ。また、V N I C 1 2 4 は、仮想MACアドレスの他に、V N I C 1 2 4 に関する情報を含んだ管理テーブル 1 5 2 を有する。管理テーブル 1 5 2 は、V N I C 1 2 4 毎に設けられる。以下、管理テーブル 1 5 2 について、L P A R # 1 1 に割り当てられたV N I C # 1 1 及び# 1 2 のそれぞれの管理テーブルを例に取る。

[0033] 図 2 は、V N I C # 1 1 の管理テーブル # 1 1 とV N I C # 1 2 の管理テーブル # 1 2 の一例を示す。

[0034] 管理テーブル # 1 1 は、V N I C # 1 1 の仮想MACアドレスと、L P A R # 1 1 のゲストOS # 1 1 が使用し得る 1 以上のMACアドレスである 1 以上のOS登録MACアドレス（すなわち、V N I C # 1 1 を経由するパケットに含まれるMACアドレスとして使用され得る 1 以上のMACアドレス）とを含む。OS登録MACアドレスは、V N I C # 1 1 の仮想MACアドレスと同じ値である場合もある。

[0035] ゲストOS # 1 1 が送信するパケットに含まれるMACアドレスが、管理テーブル # 1 1 に登録される。ハイパーバイザ 1 0 3 は、管理テーブル # 1 1 を参照することで、ゲストOS # 1 1 によって登録されたMACアドレスを知ることができる。

[0036] 例えばV N I C # 1 1 及び# 1 2 についてteamingの設定がされた場合、図 2 に示す管理テーブル # 1 1 及び# 1 2 の通り、V N I C # 1 1 とV N I C # 1 2 とでは仮想MACアドレスが異なるが、使用されるOS登録MACアドレスは、図 2 に示す通り同じになることがある（例えばMACアドレス 0 0 : 0 0 : 8 7 : 0 0 : 0 0 : 0 1 ）。teaming時には、このOS登録MACアドレスを使用すれば良いようにも思えるが、teamingの場合には、ゲストOSがハイパーバイザ 1 0 3 にどのMACアドレスを使用してくるか分からないため、登録MACアドレスを使用することができない（図 2 に示すように、両方とも共通のMACアドレスが複数登録されている場合がある）。

[0037] 再び図1を参照する。L P A R 1 2 8では、前述したように、ゲストOS 1 3 1が動作する。L P A R # 1 3のゲストOS # 1 3のようにV L A N t a gを使用しないゲストOS 1 3 1からは、図3に示すような、V L A N t a gを含まないパケットが送信される。L P A R # 1 1のゲストOS # 1 1或いはL P A R # 1 2のゲストOS # 1 2のようにt a g V L A Nを使用するゲストOS 1 3 1からは、図4に示すような、V L A N t a gを含むパケットが送信される。図3及び図4に示すいずれの種類のパケットにも、宛先M A Cアドレスと送信元M A Cアドレスが含まれる。宛先M A Cアドレスは、送信元のL P A R 1 2 8のゲストOS 1 3 1が送信先として指定するM A Cアドレスであり、送信元M A Cアドレスは、送信元のL P A R 1 2 8のゲストOS 1 3 1が送信元として指定するM A Cアドレスである。例えば、t a g V L A Nを使用しないゲストOSから送信されるパケット（以下、V L A N未使用パケット）は、第1のフィールド（例えば最初の6バイトのフィールド）に宛先M A Cアドレスを有し、第2のフィールド（例えば次の6バイトのフィールド）に送信元M A Cアドレスを有し、第3のフィールド（例えば次の2バイトのフィールド）にパケットのタイプを示す情報（例えばプロトコルを識別するための番号）を有し、第4のフィールド（例えば次の46～1500バイトのフィールド）に、データを有し、第5のフィールド（例えば最後の4バイトのフィールド）に、エラー検出のためのコードを有する。一方、t a g V L A Nを使用するゲストOSから送信されるパケット（以下、V L A N使用パケット）は、V L A N未使用パケットとほぼ同じであり、相違点は、V L A N使用パケットには、第2のフィールドと第3のフィールドとの間に第6のフィールド（例えば4バイト）がある点である。第6のフィールドには、V L A N t a gがある。具体的には、例えば、第6のフィールドの前半部分（例えば2バイト）には、T P I D（T a g Protocol Identifier）があり、第6のフィールドの後半部分（例えば2バイト）には、T C I（Tag Control Information）がある。T P I Dの値は、t a g V L A N使用時には「0x8100」となる。T C Iは、V L

AN IDやパケットの優先順位等の情報を含む情報である。TPIDとTCIとを含んだ情報を、VLAN tagと呼ぶ。

[0038] 前述したように、例えば、LPAR # 11及び# 12のようにtag VLANを使用するLPAR 128からは、VLAN ID等の情報を含んだVLAN tagを含むパケットが送信される。また、例えば、LPAR # 13のように、ネットワークを冗長化するTeamingを行うLPAR 128からは、ゲストOSがVNIC 152の仮想MACアドレスと異なるMACアドレスを含むパケットが送信されることがある。

[0039] ハイパーバイザ # 1により、稼働中のLPAR 128が物理計算機 # 1から物理計算機 # 2にLMされることがある。例えば、LPAR # 13がLMされると、LPAR # 13が物理計算機 # 2にマイグレーションされ、物理計算機 # 2内のハイパーバイザ # 2が、RARP (Reverse Address Resolution Protocol) を送信すれば、PSW # 3のMACアドレステーブル # 3とPSW # 4のMACアドレステーブル # 4が更新され、LPAR # 13のネットワーク接続を復元（実質的に引き継ぐ）ことができる。

[0040] 本実施形態では、LMされるLPAR 128がVNICの仮想MACアドレスと異なるソフトウェアMACを含んだパケットを送信するようになっていても（或いは、VLAN tagを含んだパケットを送信するようになっていても）、そのようなLPAR 128がLMされた後にネットワーク接続を復元できるよう（言い換えれば、LMに伴いネットワーク接続の引き継ぎができるよう）、少なくともLM元の物理計算機 # 1内のハイパーバイザ # 1が、LPAR 128からのパケットから所定種類の情報を抽出し保持するという学習機能を有する。所定種類の情報は、例えば、MACアドレス（送信元MACアドレス）とVLAN tagの少なくとも一部分（例えばVLAN ID）とのうちの少なくとも一方（典型的には両方）である。上記学習機能は、MAC学習部 140である。MAC学習部 140は、MACアドレス（送信元MACアドレス）とVLAN IDとを抽出し管理するMACアドレス学習を行うようになっている。

- [0041] ハイパーバイザ103は、自身が構築する仮想的なネットワークセグメント毎にVSW（仮想的なスイッチ）137を有する。VSW137は、物理的なNIC112のポートに対応している。
- [0042] 各VSW137は、複数の仮想ポートと、MAC学習部140と、MACアドレス学習テーブル144とを有する。各VSW#137の複数の仮想ポートは、LPAR128に割り当てられたVNIC124に接続される仮想ポートと、物理計算機101が有する物理的なNIC112に割り当てられる仮想ポートとを含む。MAC学習部140は、MACアドレス学習を行う。例えば、VSW#11のMAC学習部#11は、VSW#11に接続されているVNICが割り当てられているLPAR#11（ゲストOS#11）からVSW#11がパケットを受信した場合、そのパケットからMACアドレス（送信元MACアドレス）とVLAN IDとを抽出し、抽出した情報（例えばMACアドレス及びVLAN ID）をMACアドレス学習テーブル#11に登録する。これにより、MACアドレス学習テーブル140に登録される情報（MACアドレス及びVLAN ID）は、そのMACアドレス学習テーブル140を有するVSW137が受信したパケットに依存する。
- [0043] 図1に例示した構成によれば、LPAR128に割り当てられたVNIC124に少なくとも1つのVSW137が接続されており、VSW137に少なくとも1つの物理的なNIC112が接続されており、物理的なNIC112に少なくとも1つのPSW154が接続されており、PSW154に、そのPSW154に物理的なNIC112を通じて接続されているVSW（仮想ネットワークセグメント）137に対応したネットワークセグメントが接続されている。ネットワークセグメントは、例えば、LANのサブネット、或いは、VLAN（tag VLAN）で良い。例えば、ゲストOS#13から送信されたパケットは、LPAR#13に割り当てられているVNIC124、そのVNIC124に接続されているVSW#13、及び、VSW#13に接続されているNIC#2を通じて、NIC#2に接続されて

いる P S W # 3 で受信される。P S W # 3 は、その受信したパケット内の宛先 M A C アドレスに対応した物理ポート I D を M A C アドレステーブル # 3 から特定し、特定した物理ポート # A 又は # D から、その受信したパケットを出力する。これにより、パケットは、P S W # 3 に接続されているネットワークセグメント # 3 を通じて外部の装置へ送られる、或いは、物理計算機 # 2 内の L P A R 1 2 8 へと送られる。

[0044] V S W 1 3 7 は、ネットワークセグメントに対応しているので、複数の物理計算機 1 0 1 に、同一のネットワークセグメントに対応した複数の V S W 1 3 7 が存在することになる。例えば、物理計算機 # 1 及び # 2 に、同一のネットワークセグメント # 1 に対応した V S W # 1 1 と V S W # 2 1 が存在することになる。

[0045] 管理コンソール 1 4 9 は、例えば、パーソナルコンピュータのような計算機でも良い。管理コンソール 1 4 9 は、例えば、複数の P S W 1 5 4 のうちの少なくとも 1 つ（例えば P S W # 1 ）に接続されている。管理コンソール 1 4 9 から、P S W # 1 を通じて、L M 開始要求が L M 元の物理計算機 # 1 と L M 先の物理計算機 # 2 の両方に指示が出されることにより開始される。L M 開始要求は、例えば、物理計算機 # 1 （又は # 2 ）だけに送られて、その物理計算機 # 1 （又は # 2 ）から相手の物理計算機 # 2 （又は # 1 ）に L M 開始要求が送られても良い。マイグレーション処理部 # 1 が、N I C # 1 を通じてマイグレーション処理部 # 2 に L M 対象の L P A R # 1 2 8 のデータを転送することにより、L P A R 1 2 8 の L M が実現される。L M 開始要求では、L M 対象の L P A R 1 2 8 （例えば L P A R # 1 3 ）が及びマイグレーションの方式が指定される。指定された L P A R 1 2 8 が稼働中であり、かつ指定された方式が L M であれば、L M が実行される。指定された方式が L M であるにも関わらず、指定された L P A R 1 2 8 が稼働中でなければ、L M は失敗となる。本実施形態に係るマイグレーション処理部 1 4 8 は、L M の他に、他種のマイグレーション、例えば、マイグレーション対象の L P A R 1 2 8 のゲスト O S 1 3 1 をシャットダウンした状態でそのマイグレ

ーション対象のL P A R 1 2 8をマイグレーションするシャットダウンマイグレーションを行うことができる。シャットダウンマイグレーションを行う際には、マイグレーション方式としてシャットダウンマイグレーションが指定される必要がある。

[0046] L P A R 1 2 8が物理計算機# 2にL Mされた場合、L M先の物理計算機# 2のマイグレーション処理部# 2が、R A R Pパケットを送信する。

[0047] 図5は、L M対象のL P A R 1 2 8がt a g V L A Nを使用していないL P A R 1 2 8である場合に送信されるR A R Pパケットの構成例を示す。

[0048] 図5からわかるように、パケットの構成は、図3に示したパケットの構成と同じである。第3のフィールド（タイプが書き込まれるフィールド）に、R A R Pを意味する値（例えば0x8035）が書き込まれ、第4のフィールド（データを有するフィールド）に、R A R Pに関するデータ（R A R Pパケット構成に従う値）が書き込まれる。

[0049] 図6は、ハイパーバイザ103内のV S W 1 3 7が行う処理の流れを示す。

[0050] 図6の処理は、物理計算機101が起動しハイパーバイザ103によりV S W 1 3 7が構築された後、全てのV S W 1 3 7で実施される共通の処理である。具体的には、図6の処理は、V S W 1 3 7を有する物理計算機101がL P A R 1 2 8のL M元であるかL M先であるかに関係なく実施される。図6の処理は、全てのV S W 1 3 7に適用されるが、以下、本実施形態のL M対象がL P A R # 1 3であるため、V S W # 1 3を例に取り、図6に示す処理を説明する。

[0051] （S 6 0 1）物理計算機# 1に電源が入り、ハイパーバイザ# 1が動き始め、V S W # 1 3が構築され、V S W # 1 3の処理が開始する。

[0052] （S 6 0 2）V S W # 1 3は、ゲストOS# 13が送信したパケットを受信するまで、又は、学習データ取得要求を受けるまで、待機する。

[0053] （S 6 0 3）ゲストOS# 13が送信したパケットを受信した場合には、M A C学習部# 13がM A Cアドレス学習処理を行う（M A Cアドレス学習

処理の詳細については後の図9を参照)。

[0054] (S604) VSW#13は、MACアドレス学習処理を行った後、受信したパケットを送信する。そのパケットは、VSW#13に接続されているNIC#12を通じて物理計算機#1から出力される。

[0055] (S605) VSW#13は、学習データ取得要求を受信した場合には、学習データ取得処理を行う(学習データ取得処理を含む学習データ収集処理の詳細については、後の図10を参照)。

[0056] (S606) VSW#13は、学習データを取得し、取得した学習データを、学習データ取得要求の送信元(VNIC)に送信する。

[0057] 以上のように、ハイパーバイザ103が動作し、VSW137が存在しているときには、常に、各VSW137で、図6のS602~S604の処理が繰り返される。S605とS606は、LM処理が行われた場合に行われる。

[0058] 図7は、MACアドレス学習テーブル144の構成例を示す。

[0059] MACアドレス学習テーブル144は、前述したように、VSW137毎に存在し、1以上のエントリを有する。1つのエントリは、例えば、このMACアドレス学習テーブル144を有するVSW137が受信したパケットから抽出された送信元MACアドレスと、そのパケットから抽出されたVLAN IDと、そのパケットの送信元のVNICのIDと、このエントリが更新された時刻である更新時刻とを有する。すなわち、各エントリは、VNICを特定するのに、VNICの仮想MACアドレスではなく、VNICのIDを使用する。VNICについて、VNICの仮想MACアドレスが登録されても良い。

[0060] LMされたLPAR128のネットワーク接続をLM先の物理計算機#2において復元するために、ハイパーバイザ#2(マイグレーション処理部#2)がRARPパケットを出す。図5を参照して説明したが、RARPパケットには、送信元MACアドレスとVLAN ID(tag VLANが使用されていれば)が必要となる。LM元のハイパーバイザ#1におけるMA

Cアドレス学習テーブル144から、LMされたL P A Rに関する全エントリの情報をLM先のハイパーバイザ#2にマイグレーションする。これにより、ゲストOS#13がハイパーバイザ#1が管理している仮想MACアドレスと異なるMACアドレスを含んだパケットを送信していた場合（及び、t a g V L A Nが使用されていた場合）であっても、LM元から持ってきたエントリ（MACアドレス学習テーブルのエントリ）に、ネットワーク接続の復元に必要な情報であってR A R Pパケットに含められる情報が格納されている。このため、そのような情報を含んだR A R PパケットをLM先のハイパーバイザ#2（マイグレーション処理部#2）が出すことで、LMされたL P A Rのネットワーク接続を復元することができるようになる。

[0061] MACアドレス学習テーブル144は、ハイパーバイザ103によるV S W 1 3 7の初期化処理で生成され、V S W 1 3 7が存在する間、常に存在する。ゲストOS134が、生成されたどのV S W（仮想ネットワークセグメント）を使うかは、ユーザによって指定される。MACアドレス学習テーブル144の更新は、ゲストOS134が使用するパケット中の情報が更新された場合に行われる。

[0062] 図8は、メモリ111のうちハイパーバイザ103に関する情報が占める領域を示す。

[0063] 図8に示すように、MACアドレス学習テーブル144と、そのテーブル144を論理的に保持するV S W 1 3 7（正確にはV S W 1 3 7としての機能が発生させるためにC P U 1 0 4によって実行されるプログラムモジュール）とがメモリ111に配置される。これにより、ゲストOS134が大量のパケットを生成し送信しても、MACアドレス学習処理を長い時間をかけずに行うことができる。管理コンソール149のような、物理計算機#1の外部の装置がMACアドレス学習テーブル144を持っても良いかもしれないが、そうすると、通信ネットワークによりデータを転送する必要があるため、ゲストOS134により大量にパケットが生成され送信された場合に、データの解析及び更新に長い時間を要してしまうことになる。

- [0064] 図9は、図6のS603（MACアドレス学習処理）の流れを示す。
- [0065] MACアドレス学習処理は、LMとは無関係に、VSW137が存在している間、そのVSW#137のMAC学習部140で行われる処理である。ここでも、図6と同様に、VSW#13を例に取り、図9に示す処理を説明する。
- [0066] （S901）MAC学習部#13は、ゲストOS#13が送信したパケットを受信する。
- [0067] （S902）MAC学習部#13は、受信したパケットの解析を行い、ゲストOS#13がtag VLANを使用しているか否かに基づき、パケットから情報を取得する。具体的には、tag VLANが使用されている場合、取得される情報は、送信元MACアドレスとVLAN IDである。tag VLANが使用されていない場合、取得される情報は、送信元MACアドレスである。tag VLANを使用しているか否かは、例えば、パケットにおける第2のフィールド（送信元MACアドレス）の次の2バイトが、TPID「0x8100」であるか否かで判定することができる。
- [0068] （S903）MAC学習部#13は、取得した送信元MACアドレス（及びVLAN ID）と同じMACアドレス（及びVLAN ID）を含んだエントリ（図9の説明において、以下、一致エントリ）がMACアドレス学習テーブル#13にあるか否かを判定する。
- [0069] （S904）S903の判定の結果が肯定的の場合（S903：Yes）、MAC学習部#13は、一致エントリ内に、S901で受信したパケットの送信元のVNICのIDと同じVNIC IDが登録されているか否かを判定する。
- [0070] （S905）S904の判定の結果が否定的の場合（S904：No）、MAC学習部#13は、一致エントリのVNIC IDを、S901で受信したパケットの送信元のVNIC IDに変更する。これにより、例えばteaming時に異なるVNICで同じMACアドレスが使用されていた場合でも、どのVNICが使用されているかを知ることができる。

- [0071] (S906) S904の判定の結果が肯定的の場合(S904:No)、又は、S905の後、MAC学習部#13は、一致エントリにおける更新時刻を、一致エントリを更新した時刻に変更する。
- [0072] (S907) S903の判定の結果が否定的の場合(S903:No)、MAC学習部#13は、MACアドレス学習テーブル#13に空きエントリがあるか否かを判定する。
- [0073] (S908) S907の判定の結果が否定的の場合(S907:No)、MAC学習部#13は、MACアドレス学習テーブル#13から、最も更新時刻の古いエントリを削除することにより、空きエントリを生成する。
- [0074] (S909) S907の判定の結果が肯定的の場合(S907:No)、又は、S908の後、MAC学習部#13は、空きエントリに、S902で取得した情報(学習した情報)と、S901で受信したパケットの送信元のVNIC IDと、空きエントリに情報を格納する時刻である更新時刻とを、格納する。
- [0075] 図10は、学習データ収集処理の流れを示す。
- [0076] 学習データ収集処理を説明する際の前提として、図7で示したように、MACアドレス学習テーブル144は、複数のエントリを有し、各エントリは、VNIC IDを含んでいる。VNIC IDは、名前の通り、VNICを識別するための情報であり、VNICと1:1で対応する。
- [0077] LMされたLPAR128のネットワーク接続を復元するためには、LM対象のLPARに割り当てられたVNICに対応するVNIC IDを持つエントリの情報を取得し、LM先のハイパーバイザ#2が、当該エントリのMACアドレスとVLAN IDを利用してRARPパケットを送信する必要がある。LPARに割り当てられたVNICのエントリをMACアドレス学習テーブルから検索する際には、VNIC IDをキーとして検索する。LPARに複数のVNICが割り当てられていれば、複数のエントリの情報を取得する必要がある。また、LM対象のLPAR128がtag VLANを使用するLPARであることもあるので、その際には、1つのVNIC

について複数のエントリの情報を取得する必要がある。以下、図6及び図9と同様、LM対象のL P A R 1 2 8がL P A R # 1 3であり物理計算機# 1から物理計算機# 2にLMされることを例に取り、図10に示す学習データ収集処理を説明する。

[0078] (S 1 0 0 1) マイグレーション処理部# 1は、L P A R # 1 3に割り当てられている1以上のV N I C 1 3 7のうちの1つのV N I C 1 3 7（例えば# 1 1）に、学習データ取得要求を出す。

[0079] (S 1 0 0 2) V N I C # 1 1は、S 1 0 0 1の学習データ取得要求を受信し、V N I C # 1 1に接続されているV S W # 1 3に対して、学習データ取得要求を送信する。

[0080] (S 1 0 0 3) V S W # 1 3が、データ収集要求を受信する。

[0081] (S 1 0 0 4) V S W # 1 3が、V N I C # 1 1のV N I C I Dを特定する。

[0082] (S 1 0 0 5) V S W # 1 3が、M A Cアドレス学習テーブル# 1 3内に、S 1 0 0 4で特定したV N I C I Dと同一のV N I C I Dを持つエントリ（図10の説明において、以下、一致エントリ）の検索を行う。

[0083] (S 1 0 0 6) V S W # 1 3が、検索したエントリが一致エントリであるかどうかの判定を行う。

[0084] (S 1 0 0 7) S 1 0 0 6の判定の結果が肯定的の場合（S 1 0 0 6 : Y e s）、V S W # 1 3が、一致エントリの情報を、マイグレーション処理部# 1のバッファ領域（例えばメモリ11における所定領域）にコピーする。これにより、マイグレーション処理部# 1がLM先にL P A R # 1 3に関するデータを転送できるようにする。

[0085] (S 1 0 0 8) S 1 0 0 6の判定の結果が否定的の場合（S 1 0 0 6 : N o）、或いは、S 1 0 0 7の後、V S W # 1 3が、検索するエントリが最終エントリであるか否かを判定する。この判定の結果が否定的の場合（S 1 0 0 8 : N o）、別の未検索のエントリについて、S 1 0 0 5が行われる。

[0086] (S 1 0 0 9) S 1 0 0 8の判定の結果が肯定的の場合（S 1 0 0 8 : Y

e s)、VSW#13が、S1003で受信した要求の送信元のVNIC#11に対して、データ取得の完了を通知する。

[0087] (S1010) データ取得の完了の通知を受けたVNIC#11は、S1001の要求の送信元のマイグレーション処理部#1に対して、データ取得の完了を通知する。

[0088] (S1011) VNIC#11からデータ取得の完了の通知を受けたマイグレーション処理部#1は、LM対象のLPAR#13に割り当てられている1以上のVNIC137の全てについて、データの取得が完了したか否かを判定する。

[0089] (S1012) S1011の判定の結果が否定的の場合(S1011:No)、マイグレーション処理部#1は、LPAR#13に割り当てられている1以上のVNIC137のうち別のVNIC(例えば#12)に対して、S1001のように学習データ取得要求を送信する。

[0090] (S1013) S1011の判定の結果が肯定的の場合(S1011:Yes)、マイグレーション処理部#1は、学習データ収集処理を終了する。

[0091] 図11は、LM処理の流れを示す。なお、これまで、LPAR128が物理計算機#1から物理計算機#2にLMされると説明したが、正確には、物理計算機#1におけるLM元のLPARに関するデータが物理計算機#1から物理計算機#2に転送され、物理計算機#2におけるLM先のLPARに関するデータが、物理計算機#1におけるLM元のLPARに関するデータに変更されることである。そこで、以下の説明では、「LM元LPAR」及び「LM先LPAR」という言葉が使用される。このLM処理では、LM元LPARからLM先LPARにLPAR#3がマイグレーションされるというイメージである。

[0092] (S1101) 管理コンソール149が、LM元の物理計算機#1におけるマイグレーション処理部#1に対してLM開始要求を出す。

[0093] (S1102) 管理コンソール149が、LM先の物理計算機#1におけるマイグレーション処理部#2に対してLM開始要求を出す。

- [0094] (S 1 1 0 3) マイグレーション処理部 # 1 は、L M元の物理計算機 # 1 に関する構成情報を収集する。構成情報は、例えば、どのN I C 1 1 2のポートにどのV S W 1 3 7を通じてどのV N I C 1 2 4が接続されているか、V N I C毎の仮想M A Cアドレス、及び、V H B A毎のWWN (World Wide Name) を含んでよい。
- [0095] (S 1 1 0 4) マイグレーション処理部 # 2 はL M先の物理計算機 # 2 に関する構成情報を収集する。
- [0096] (S 1 1 0 5) マイグレーション処理部 # 2 は、S 1 1 0 4 で取得した構成情報のうちの、L M先L P A RのM A Cアドレス及び／又はWWNを、マイグレーション処理部 # 1 に転送する。
- [0097] (S 1 1 0 6) マイグレーション処理部 # 1 は、S 1 1 0 3 で取得した構成情報のうちの、L M元L P A RのM A Cアドレス及び／又はWWNを、マイグレーション処理部 # 2 に転送する。
- [0098] (S 1 1 0 7) マイグレーション処理部 # 2 は、S 1 1 0 6 でマイグレーション処理部 # 1 から送信されたM A Cアドレス及び／又はWWNを受信し、それらが、L M先L P A RのM A Cアドレス及び／又はWWNを、受信したM A Cアドレス及び／又はWWNに変更する。
- [0099] (S 1 1 0 8) マイグレーション処理部 # 1 は、マイグレーション処理部 # 2 に対して、リソース確保要求を出す。
- [0100] (S 1 1 0 9) マイグレーション処理部 # 2 は、マイグレーション処理部 # 1 からリソース確保要求を受信し、L M先L P A Rのリソース確保処理を行う。
- [0101] (S 1 1 1 0) リソースを確保したマイグレーション処理部 # 2 は、マイグレーション処理部 # 1 からV M E M 1 2 3 に記憶されている情報が転送されて来るのを待つ。
- [0102] (S 1 1 1 1) マイグレーション処理部 # 1 は、V M E M 1 2 3 から、L M元L P A Rに関する情報を取得し、取得した情報をマイグレーション処理部 # 2 に転送する。マイグレーション処理部 # 1 は、V M E M 1 2 3 に記憶

されている、LM元LPARに関する情報のうち、ダーティ比率が一定値以下になるまで、LM元LPARに関する情報の転送を続ける。なお、ダーティ比率とは、LM元LPARに関する情報の量に対する、LM元LPARに関する情報のうちのダーティ情報の量の比率である。ダーティ情報は、本LM処理において転送された後に更新され転送が必要になった情報である。

[0103] (S 1 1 1 2) マイグレーション処理部# 2は、ハイパーバイザ# 2におけるVMEMにおける情報のうち、LM先LPARに関する情報を、マイグレーション処理部# 1より転送されて来た情報(LM元LPARに関する情報)に変更する。

[0104] (S 1 1 1 3) マイグレーション処理部# 1は、ダーティ比率が一定値以下になった場合に、LM元LPARを休止する。具体的には、マイグレーション処理部# 1は、LM元LPARに割り当てられていた仮想リソース(VCPU 1 1 9、VNIC 1 2 4等)を休止させることにより、LM元LPARを休止させる。これにより、VMEMの情報のうちの、LM元LPARに関する情報を確定することができ、以って、ダーティ情報が増えないようにすることができる。

[0105] (S 1 1 1 4) LM元LPARの休止後に、学習データ収集処理(図10参照)が行われる。その処理は、マイグレーション処理部# 1がLM元LPARに接続されているVNICに学習データ取得要求を送信することを含む。この処理により、LM元LPARについて学習されたデータ(ソフトウェアMAC、VLAN ID)を、マイグレーション処理部# 1が取得することができる。具体的には、図10のS 1 0 0 7の通り、MACアドレス学習テーブル1 4 4の1以上のエントリの情報が、マイグレーション処理部# 1のバッファ領域にコピーされる。

[0106] (S 1 1 1 5) マイグレーション処理部# 1は、マイグレーション処理部# 1のバッファ領域にコピーされた情報(MACアドレス学習テーブル1 4 4の1以上のエントリの情報)を含んだ情報であるデバイス情報をマイグレーション処理部# 2に転送する。

- [0107] (S 1 1 1 6) マイグレーション処理部 # 2 は、S 1 1 1 5 でマイグレーション処理部 # 1 から転送されたデバイス情報を受信し、受信したデバイス情報を、LM先LPARに反映する。具体的には、マイグレーション処理部 # 2 は、LM先LPARのVNICに接続されているVSWのMACアドレス学習テーブルにおける、LM先LPARのVNICのIDに対応した1以上のエントリに、それぞれ、受信したデバイス情報のうちのVNIC ID以外の情報(MACアドレス、VLAN ID及び更新時刻)を格納する。
- [0108] ここまでの処理により、LMの前後、例えば、S 1 1 1 3の前とS 1 1 1 6の後で、図12及び図13に示す変化が生じる。
- [0109] すなわち、例えば、図12の左側に示すように、LM元LPARに割り当てられていたVNICのIDが「1」であった場合、LM後に、VNIC ID「1」を含んだ全てのエントリは、図12の右側に示すように、空きエントリとなる。一方、図13の左側に示すように、LM先LPARに割り当てられていたVNIC(ID「9」のVNIC)に対応するエントリに、受信したデバイス情報のうちのVNIC ID「1」以外の情報(MACアドレス、VLAN ID及び更新時刻)が格納される。格納先のエントリでは、VNIC IDは「9」である。
- [0110] (S 1 1 1 7) マイグレーション処理部 # 1 は、デバイス情報の転送後、最終メモリ転送を行う。例えば、LM元LPARについての残りの全てのダーティ情報がマイグレーション処理部 # 2 に転送される。
- [0111] (S 1 1 1 8) マイグレーション処理部 # 2 は、最終メモリの反映を行う。例えば、マイグレーション処理部 # 2 は、受信したダーティ情報を、ハイパーバイザ # 2 におけるVMEMに格納する。
- [0112] (S 1 1 1 9) マイグレーション処理部 # 1 は、LMされたLPARを指定したLPAR再開要求をマイグレーション処理部 # 2 に出す。
- [0113] (S 1 1 2 0) マイグレーション処理部 # 2 は、LPAR再開要求を受信し、その要求に従い、LM元で停止させていたゲストOS 1 3 1 を再開させるためのLPAR再開処理を行う。

- [0114] (S 1 1 2 1) マイグレーション処理部 # 1 は、LM元LPARのdeactivate処理を行う。
- [0115] (S 1 1 2 2) マイグレーション処理部 # 2 は、LM先LPARに接続されたVSWのMACアドレス学習テーブルから、LM先LPARに割り当てられているVNICに対応するエントリを見つけ出し、見つかったエントリ内のMACアドレス（及びVLAN ID）を含んだRARPパケットをブロードキャスト送信する。これにより、図14及び図15に示すように、そのRARPパケット内のMACアドレス（及びVLAN ID）と同じMACアドレス（及びVLAN ID）を含んだMACアドレステーブル # 1 を有するPSW # 1 が、そのRARPパケットを受信した場合に、そのMACアドレスに対応する物理ポートのID「port B」を、今回そのRARPパケットを受信した物理ポートのID「port D」に変更する。すなわち、RARPパケットのブロードキャスト送信により、PSW154に、ネットワーク経路が変わったことを伝えることができる。なお、本実施形態では、PSW154のMACアドレステーブル148を変更するためにRARPパケットのブロードキャスト送信が利用されるが、RARPパケット以外の種類のパケットにより、PSW154のMACアドレステーブル148が変更されても良い。
- [0116] (S 1 1 2 3) マイグレーション処理部 # 2 は、LM先LPARの構成情報を保存する。LPARの構成情報は、メモリに保存されても良いが、本実施形態では、物理計算機101の保守又は管理のための計算機であるSVP（Service Processor）に保存される。SVPは、複数の物理計算機101に共通であっても良いし、物理計算機101毎に用意されても良い。例えば、S1123では、LPARの構成情報は、物理計算機 # 2 のSVPに保存される。
- [0117] (S 1 1 2 4) マイグレーション処理部 # 2 は、マイグレーション処理部 # 1 に対して構成情報保存完了報告を送信する。
- [0118] (S 1 1 2 5) マイグレーション処理部 # 1 は、構成情報保存完了報告を

受信し、S 1 1 0 5でマイグレーション処理部# 2から送信されたMACアドレス及び／又はWWNを、LM元LPARのMACアドレス及び／又はWWNとして設定する。

[0119] (S 1 1 2 6) マイグレーション処理部# 1は、LM元LPARの定義（例えば、LM元LPARに関する、リソース管理部# 1内の情報）を削除する。

[0120] (S 1 1 2 7) マイグレーション処理部# 1は、マイグレーション処理部# 2に対してLPAR定義取り消し報告を送信する。

[0121] (S 1 1 2 8) マイグレーション処理部# 1は、LM元LPARの構成情報を保存する。例えば、S 1 1 2 8では、LPARの構成情報は、物理計算機# 1のSVPに保存される。

[0122] (S 1 1 2 9) マイグレーション処理部# 1は、構成情報保存完了後、LM先マイグレーション処理部172に対して構成情報保存完了報告を送信する。

[0123] (S 1 1 3 0) マイグレーション処理部# 1は、マイグレーション処理部# 2に構成情報保存完了報告を送信した後、管理コンソール149に対してLM完了報告を送信する。

[0124] (S 1 1 3 1) マイグレーション処理部# 2は、マイグレーション処理部# 1から構成情報保存完了報告を受信し、管理コンソール149に対してLM完了報告を送信する。

[0125] (S 1 1 3 2) 管理コンソール149は、マイグレーション処理部# 1及び# 2の両方からLM完了報告を受信した後、LMを完了させる。

[0126] 上述した実施形態によれば、ハイパーバイザ103の管轄外のMACアドレス或いはVLAN IDを使用した通信が行われる環境であっても、管理者によって当該MACアドレス或いはVLAN IDが手動で入力されること無しに、ライブマグレーションされたLPARのネットワーク接続を復元することができる。管理者が手動でハイパーバイザ103の管轄外のMACアドレス或いはVLAN IDを入力する必要が無いので、当該情報入力の

ためにL P A R 或いはその上位マシンの使用者とハイパーバイザ 1 0 3 の管理者が連携する必要も無い。

[0127] また、複数のネットワークセグメントにそれぞれ対応した複数のV S W 1 3 7 がハイパーバイザ 1 0 3 で構築される場合、V S W 1 3 7 毎にM A C アドレス管理テーブル 1 4 4 がある。故に、ハイパーバイザ # 1 は、複数のM A C アドレス管理テーブル 1 4 4 を管理する。このような構成であっても、L M 対象のL P A R 1 2 8 に割り当てられているV N I C を通じてV S W から学習データを収集することにより、L M 対象のL P A R 1 2 8 ではないL P A R 1 2 8 に関するデータをL M 先に転送することを回避することができる。

[0128] 以上、一実施形態を説明したが、本発明は、その実施形態に限定されるものでなく、その趣旨を逸脱しない範囲で種々変更可能であることはいうまでもない。例えば、上述の実施形態では、V N I C を介した通信を例に取ったが、他種の仮想 I / O デバイス（例えばV H B A）でも同様の処理が行われても良い。また、ハイパーバイザ 1 0 3 のような仮想化機構は、コンピュータプログラムとして実現されても良いし、物として実現されても良い。例えば、ハイパーバイザは、コンピュータプログラムを記憶した記憶デバイスと、物理計算機のような外部装置と通信するための通信インタフェースデバイスと、記憶デバイス及び通信インタフェースデバイスに接続されておりプロセッサを含んだ制御デバイスを有した装置であって、制御デバイス内のプロセッサでコンピュータプログラムが実行されることにより、ハイパーバイザとしての機能が実現されて良い。また、例えば、ハイパーバイザ 1 0 3 が有するスケジューラ 1 1 7、リソース管理部 1 1 8 及びマイグレーション処理部 1 4 8 のうちの少なくとも 1 つがハードウェア回路で実現されても良いし、それらの要素 1 1 7、1 1 8 及び 1 4 8 の少なくとも 1 つの一部の処理がハードウェア回路で実現されても良い。

符号の説明

[0129] 1 0 1 : 物理計算機、1 0 3 : ハイパーバイザ、1 2 8 : L P A R（仮想計

算機)、137:VSW(仮想スイッチ)、140:MAC学習部、144
:MAC学習テーブル

請求の範囲

[請求項1]

1以上の物理ネットワークスイッチに接続される1以上の第1の物理I/O（Input/Output）デバイスを含む第1の物理リソースと、前記第1の物理リソースを論理的に分割することにより得られた1以上の第1の仮想リソースをそれぞれ1以上の第1のL P A R（Logical Partition）に割り当てることで1以上の仮想計算機を構築することができる第1の仮想化機構と、を有する第1物理計算機と、

前記1以上の物理ネットワークスイッチに接続される1以上の第2の物理I/Oデバイスを含む第2の物理リソースと、前記第2の物理リソースを論理的に分割することにより得られた1以上の第2の仮想リソースをそれぞれ1以上の第2のL P A Rに割り当てることで1以上の仮想計算機を構築することができる第2の仮想化機構と、を有する第2の物理計算機とを有し、

（A）前記第1の仮想化機構は、第1の管理情報ユニットを有し、
仮想計算機からパケットを受信した場合、前記パケットから、所定種類の情報を抽出し、前記抽出した所定種類の情報を前記第1の管理情報ユニットに登録し、

（B）前記第1の仮想化機構は、稼働中のマイグレーション対象の仮想計算機である対象仮想計算機を前記第1の物理計算機から前記第2の物理計算機にマイグレーションするライブマイグレーションのときに、前記第1の管理情報ユニット中の前記所定種類の情報を、前記第2の仮想化機構に送信し、

（C）前記第2の仮想化機構は、第2の管理情報ユニットを有し、
前記所定種類の情報を前記第2の仮想化機構から受信し、
前記受信した所定種類の情報を前記第2の管理情報ユニットに登録し、

前記所定種類の情報を含んだ所定種類のパケットを、前記1以上

の第2の物理 I / O デバイスの少なくとも1つから送信する、
仮想計算機システム。

[請求項2]

前記第1の仮想リソースは、前記1以上の第1の物理 I / O に基づく第1の仮想 I / O デバイスを含み、前記1以上の第2の仮想リソースは、前記1以上の第2の物理 I / O に基づく第2の仮想 I / O デバイスを含み、

前記第1の仮想化機構は、前記第1の仮想化機構が管理する1以上の仮想計算機にそれぞれ割り当てられる1以上の第1の仮想 I / O デバイスと前記1以上の第1の物理 I / O デバイスの少なくとも1つとが接続される1以上の第1の仮想スイッチを有し、

前記第2の仮想化機構は、前記第2の仮想化機構が管理する1以上の仮想計算機にそれぞれ割り当てられる1以上の第2の仮想 I / O デバイスと前記1以上の第2の物理 I / O デバイスの少なくとも1つとが接続される1以上の第2の仮想スイッチを有し、

各第1の仮想スイッチが、前記第1の管理情報ユニットを有し、各第2の仮想スイッチが、前記第2の管理情報ユニットを有し、

前記(A)に関して、前記第1の仮想化機構が、仮想計算機の第1の仮想 I / O デバイスを通じてその仮想計算機からパケットを受信した第1の仮想 I / O デバイスの第1の管理情報ユニットに、前記受信したパケットに含まれている所定種類の情報を登録し、

前記(B)に関して、前記第1の仮想化機構は、前記ライブマイグレーションのときに、前記対象仮想計算機の第1の仮想 I / O デバイスから出力されたパケットが経由する第1の仮想スイッチが有する第1の管理情報ユニット中の少なくとも前記対象仮想計算機の第1の仮想 I / O デバイスに対応し前記所定種類の情報を含んだ対象情報部分を、前記第2の仮想化機構に送信し、

前記(C)に関して、前記第2の仮想化機構が、マイグレーションされた対象仮想計算機の第2の仮想 I / O デバイスに接続されている

第2の仮想スイッチの第2の管理情報ユニットに、前記対象情報部分を登録し、前記対象情報部分中の前記所定種類の情報を含んだ前記所定種類のパケットを、前記対象仮想計算機の第2の仮想I/Oデバイスが接続された第2の仮想スイッチに接続されている第2の物理I/Oデバイスから送信する、
請求項1記載の仮想計算機システム。

[請求項3] 各管理情報ユニットには、それぞれ、当該管理情報ユニットを有する仮想スイッチに接続されている仮想I/OデバイスのデバイスIDと、その仮想I/Oデバイスを通じて受信したパケット中の前記所定種類の情報とが登録される、
請求項2記載の仮想計算機システム。

[請求項4] 前記(A)に関して、前記抽出した所定種類の情報と同一の情報が、前記パケットを受けた第1の仮想スイッチの第1の管理情報ユニットに登録されているが、その登録済の情報に対応したデバイスIDから識別される仮想I/Oデバイスが、前記パケットの送信元の第1の仮想I/Oデバイスの異なっている場合、前記第1の仮想化機構が、前記パケットを受けた第1の仮想スイッチの第1の管理情報ユニットにおける前記登録済の情報に対応したデバイスIDを、前記パケットの送信元の第1の仮想I/OデバイスのIDに更新する、
請求項3記載の仮想計算機システム。

[請求項5] 前記所定種類の情報は、MACアドレス及びVLAN (Virtual LAN) tagのうちの少なくとも1つである、
請求項1記載の仮想計算機システム。

[請求項6] 前記所定種類の情報は、MACアドレスであり、
前記所定種類のパケットは、RARP (Reverse Address Resolution Protocol) 要求であり、
前記RARPパケットは、MACアドレスを含み、
各物理ネットワークスイッチは、物理ポートIDとMACアドレス

の対応関係を表すアドレス管理情報とを有し、パケットを受信した場合、対象MACアドレスが前記アドレス管理情報に無ければ、前記対象MACアドレスと対象物理ポートIDとの組を前記アドレス管理情報に登録し、前記対象MACアドレスが前記アドレス管理情報に有り、且つ、前記対象MACアドレスに対応している物理ポートIDが、前記対象物理ポートIDと異なっていれば、前記アドレス管理情報における、前記対象MACアドレスに対応した物理ポートIDを前記対象物理ポートIDに更新するようになっており、

前記対象MACアドレスは、前記受信したパケットが有するMACアドレスであり、

前記対象物理ポートは、前記パケットを受信した物理ポートである、
請求項5記載の仮想計算機システム。

[請求項7] 前記所定種類の情報及び前記RARPパケットは、MACアドレスの他にVLAN tagで含んでおり、

前記アドレス管理情報は、MACアドレスと物理ポートIDとVLAN tag部分の少なくとも一部分との対応関係を表しており、

各物理ネットワークスイッチは、パケットを受信した場合、前記対象MACアドレスが前記アドレス管理情報に無ければ、前記対象MACアドレスと対象VLAN tag部分と前記対象物理ポートIDとの組を前記アドレス管理情報に登録し、前記対象MACアドレスが前記アドレス管理情報に有り、且つ、前記対象MACアドレスに対応しているVLAN tag部分が、前記対象VLAN tag部分と異なっていれば、前記アドレス管理情報における新たなエントリに、前記対象MACアドレスと前記対象VLAN tag部分と前記対象物理ポートIDとの組を登録するようになっており、

前記対象VLAN tag部分は、前記受信したパケットが有するVLAN tagの少なくとも一部分である、

請求項 6 記載の仮想計算機システム。

[請求項 8] 前記 (B) に関して、前記第 1 の仮想化機構は、前記対象仮想計算機に割り当てられている仮想リソースを休止し、その後に、前記第 1 の管理情報ユニット中の前記所定種類の情報を、前記第 2 の仮想化機構に送信する、

請求項 1 記載の仮想計算機システム。

[請求項 9] 1 以上の物理ネットワークスイッチに接続される 1 以上の物理 I/O (Input/Output) デバイスを有する物理計算機で実行され、前記物理リソースを論理的に分割することにより得られた 1 以上の仮想リソースをそれぞれ 1 以上の L P A R (Logical Partition) に割り当てることで 1 以上の仮想計算機を構築することができる仮想化機構であって、

稼働中のマイグレーション対象の仮想計算機である対象仮想計算機を物理計算機から別の物理計算機にマイグレーションするライブマイグレーションを実行することができるマイグレーション処理部と、

管理情報ユニットと、

仮想計算機からのパケットから、所定種類の情報を抽出し、前記抽出した所定種類の情報を前記管理情報ユニットに登録する学習部とを有し、

前記ライブマイグレーションのときに、前記マイグレーション処理部は、前記管理情報ユニット中の前記所定種類の情報を取得し、取得した前記所定種類の情報を、前記別の物理計算機に送信する、仮想化機構。

[請求項 10] 1 以上の物理ネットワークスイッチに接続される 1 以上の物理 I/O (Input/Output) デバイスとプロセッサと記憶デバイスとを含む物理リソースを有し、

前記プロセッサが、仮想化機構を実行することで、前記物理リソースを論理的に分割することにより得られた 1 以上の仮想リソースをそ

れぞれ 1 以上の L P A R (Logical Partition) に割り当てることで 1 以上の仮想計算機を構築することができ、且つ、稼働中のマイグレーション対象の仮想計算機である対象仮想計算機を物理計算機から別の物理計算機にマイグレーションするライブマイグレーションを実行することができるようになっており、

前記記憶デバイスは、管理情報ユニットを記憶し、

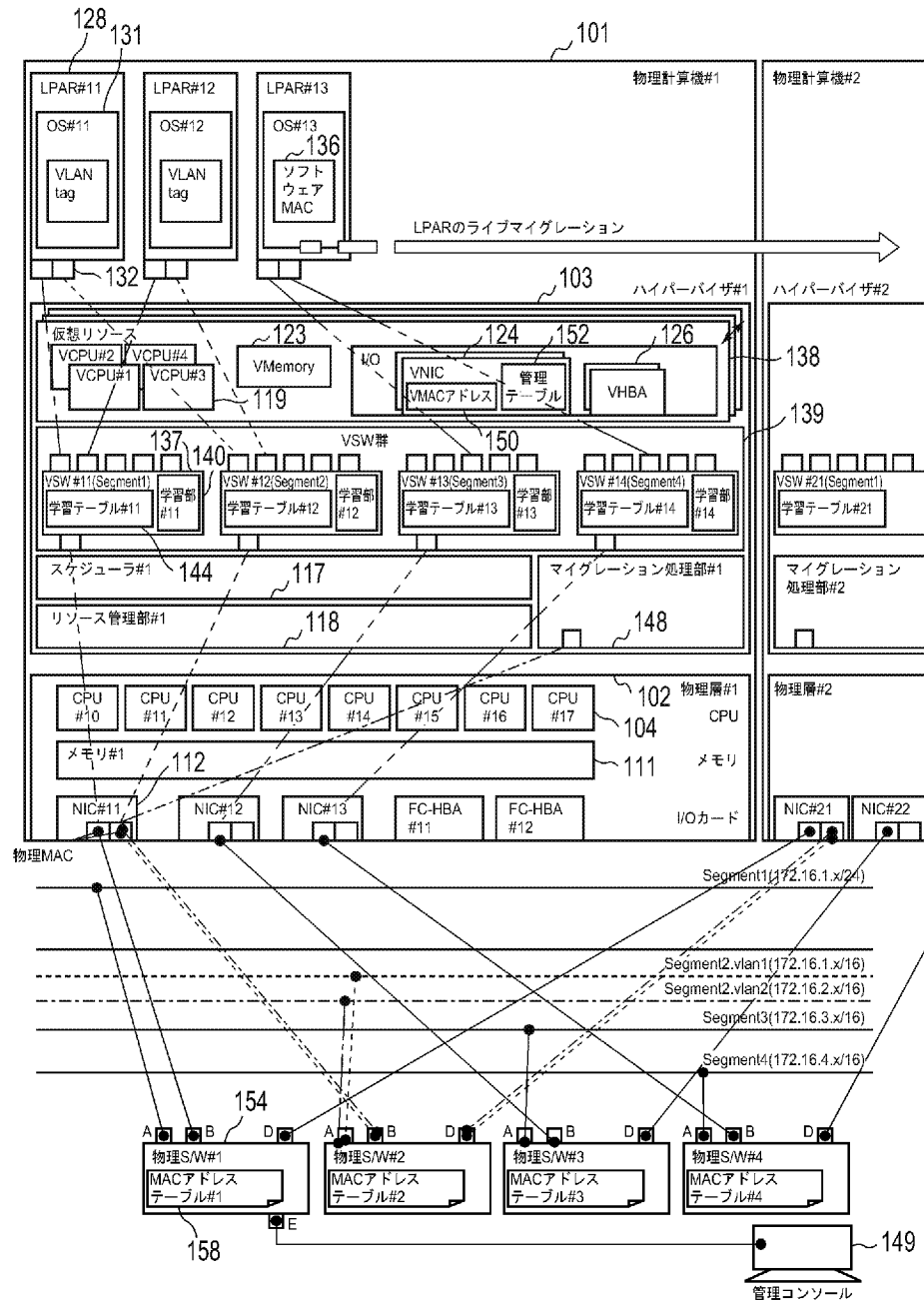
前記プロセッサは、仮想計算機からのパケットから、所定種類の情報を抽出し、前記抽出した所定種類の情報を前記管理情報ユニットに登録し、

前記ライブマイグレーションのときに、前記プロセッサは、前記管理情報ユニット中の前記所定種類の情報を、前記別の物理計算機に送信する、

物理計算機。

[図1]

Fig. 1



[図2]

Fig. 2

LPAR#1のVNIC#11が有する管理テーブル#11			LPAR#1のVNIC#12が有する管理テーブル#12		
VMACアドレス		00:00:87:00:00:01	VMACアドレス		00:00:87:00:00:02
OS登録MACアドレス1		00:00:87:00:00:01	OS登録MACアドレス1		00:00:87:00:00:01
.	.	.	.	.	.
OS登録MACアドレスn		00:00:87:00:00:05	OS登録MACアドレスn		00:00:87:00:00:06

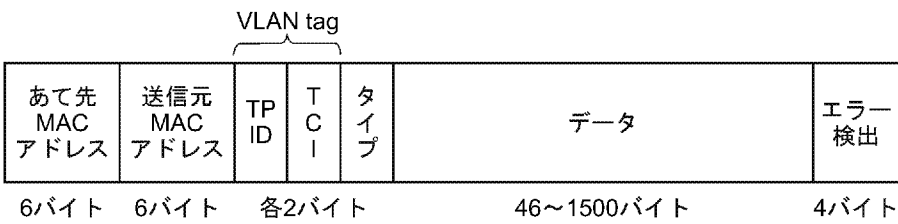
[図3]

Fig. 3

あて先 MAC アドレス	送信元 MAC アドレス	タイ プ	データ	エラー 検出
6バイト	6バイト	2バイト	46～1500バイト	4バイト

[図4]

Fig. 4



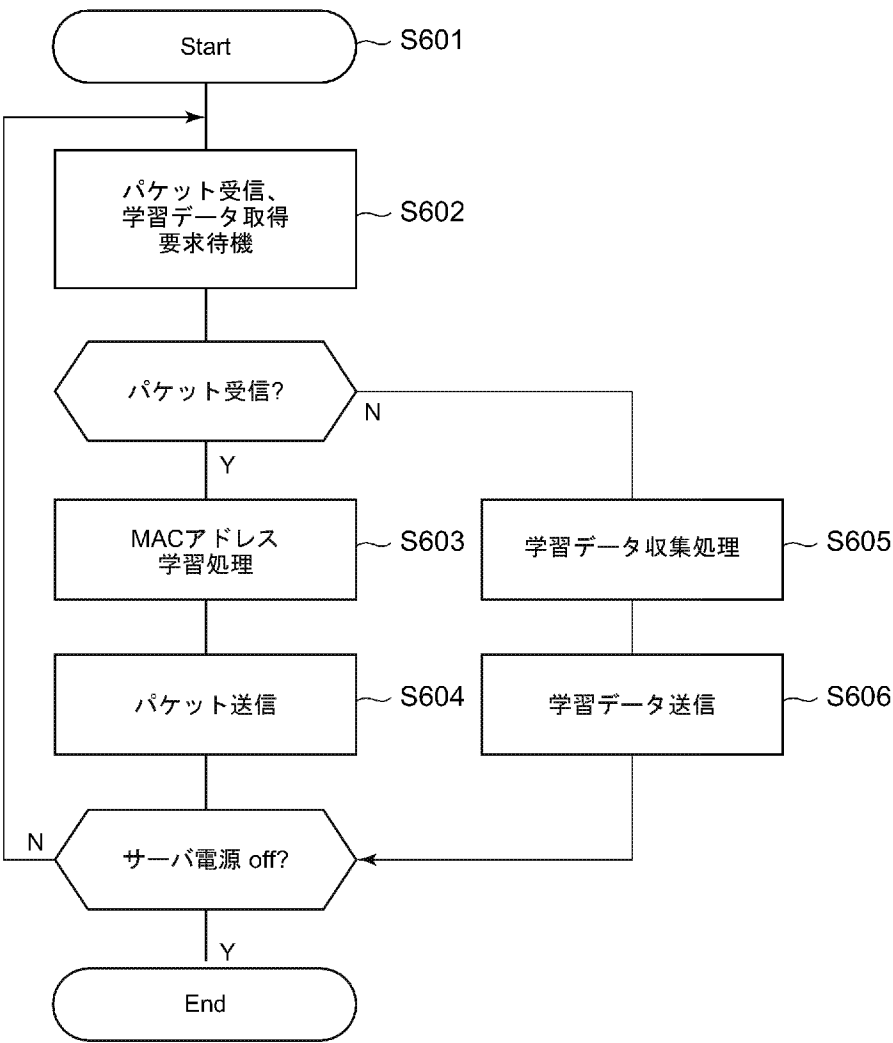
[図5]

Fig. 5



[図6]

Fig. 6



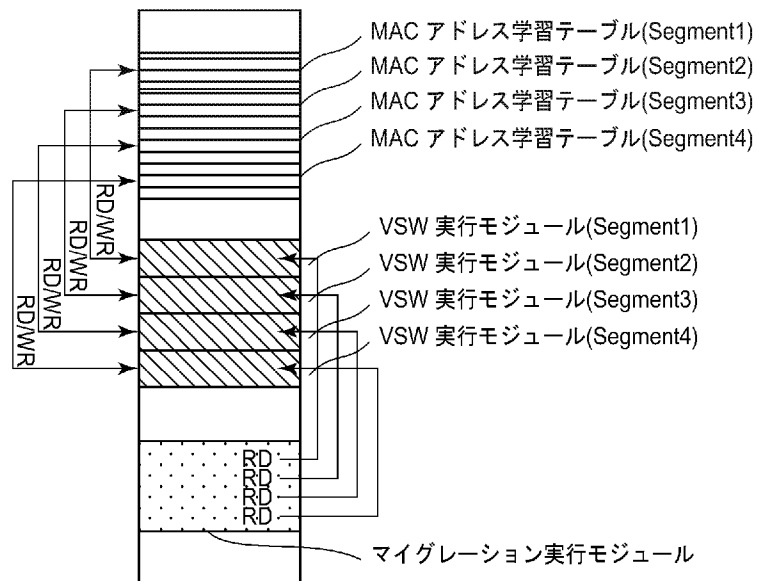
[図7]

Fig. 7

#	内容	具体例
1	MAC Address	00:00:87:00:00:01
2	VLAN ID	1
3	VNIC ID	1
4	更新時刻	2012/06/28 10時0分0秒

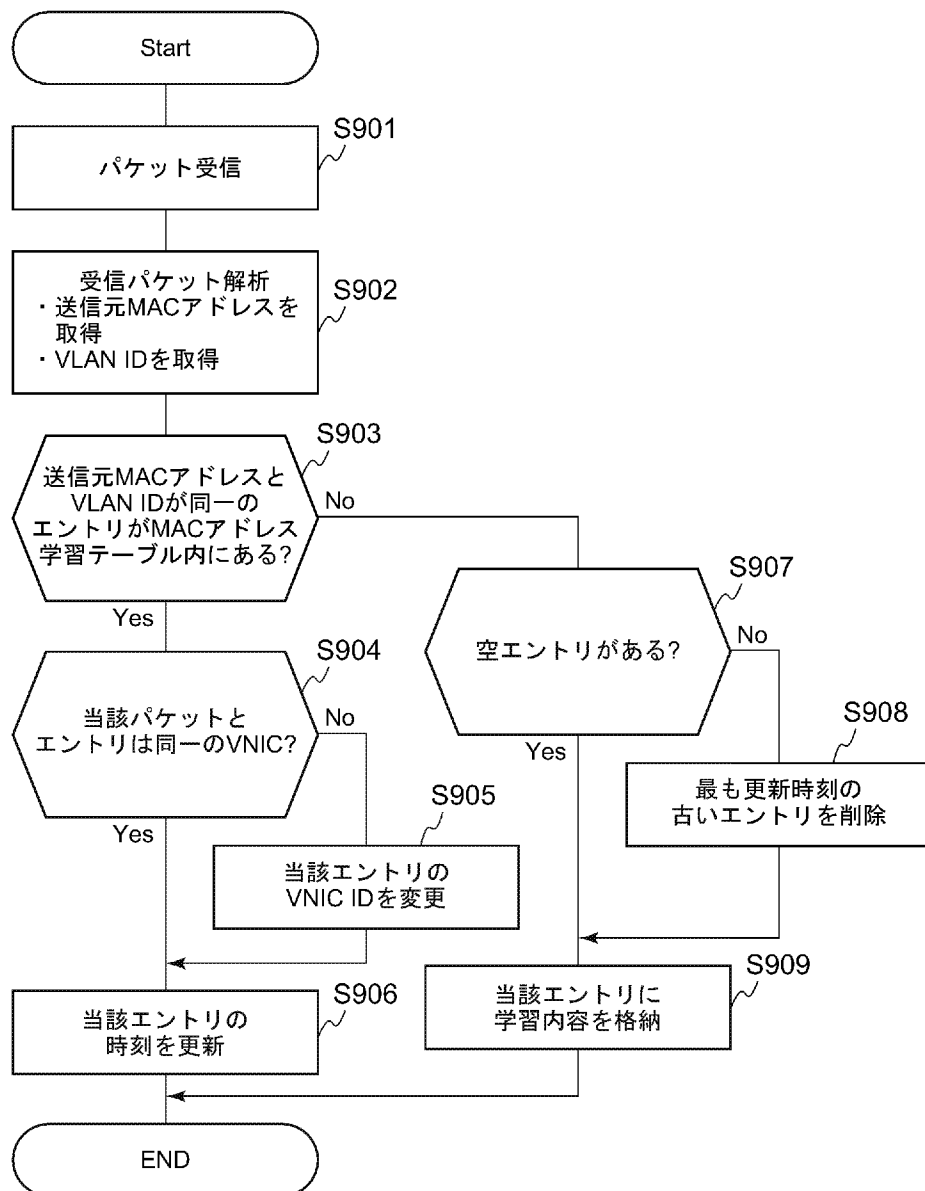
[図8]

Fig. 8



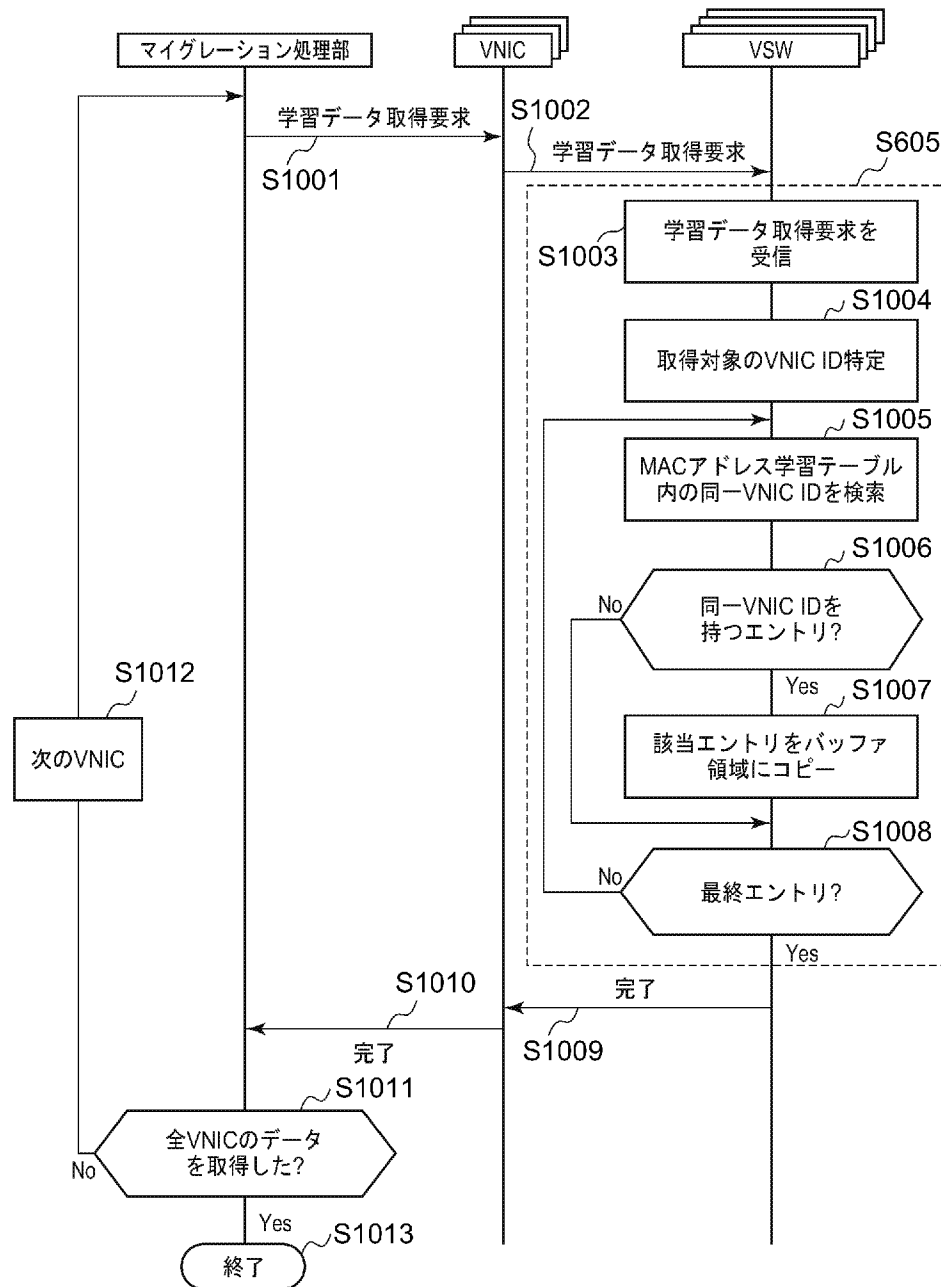
[図9]

Fig. 9



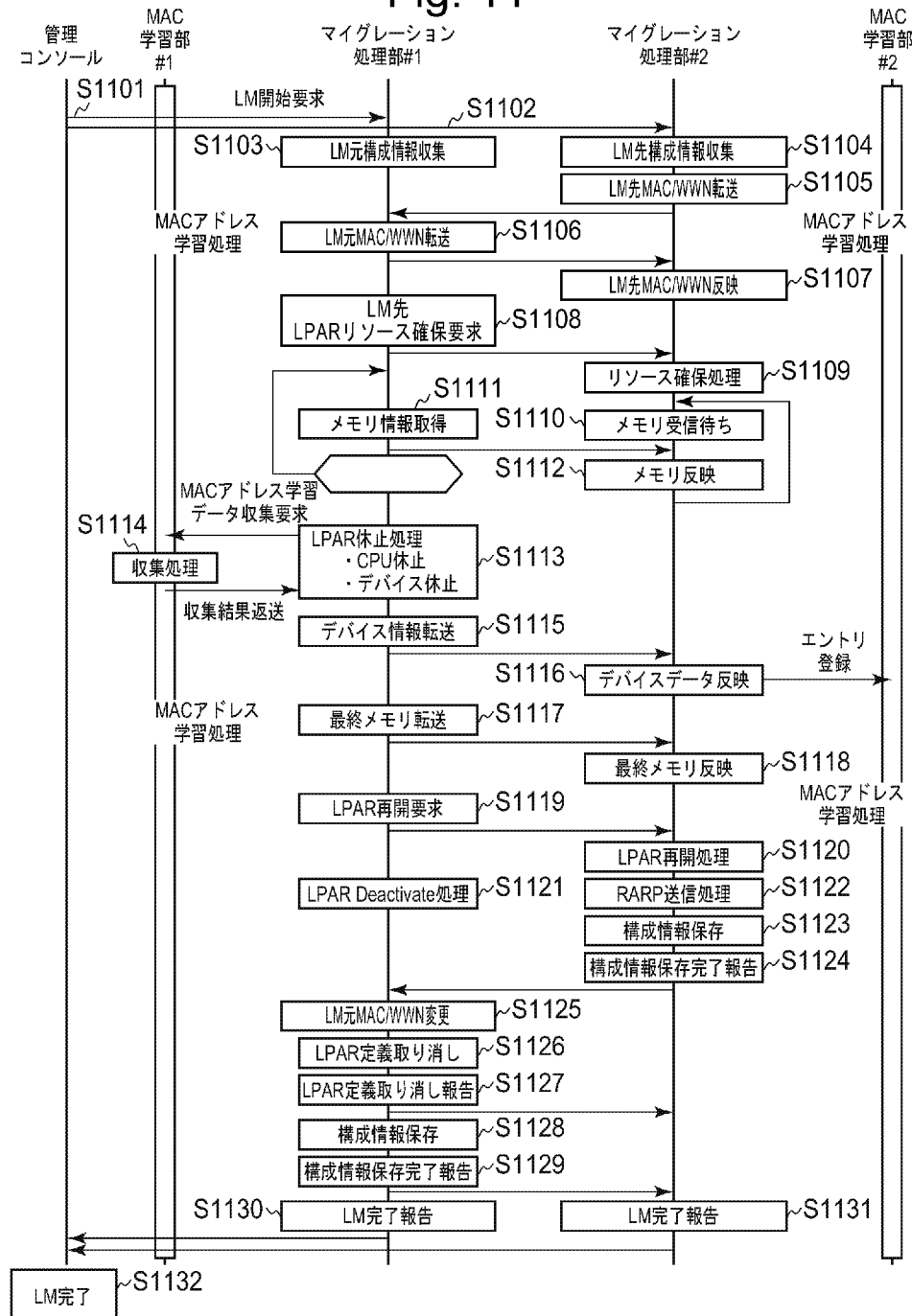
[図10]

Fig. 10



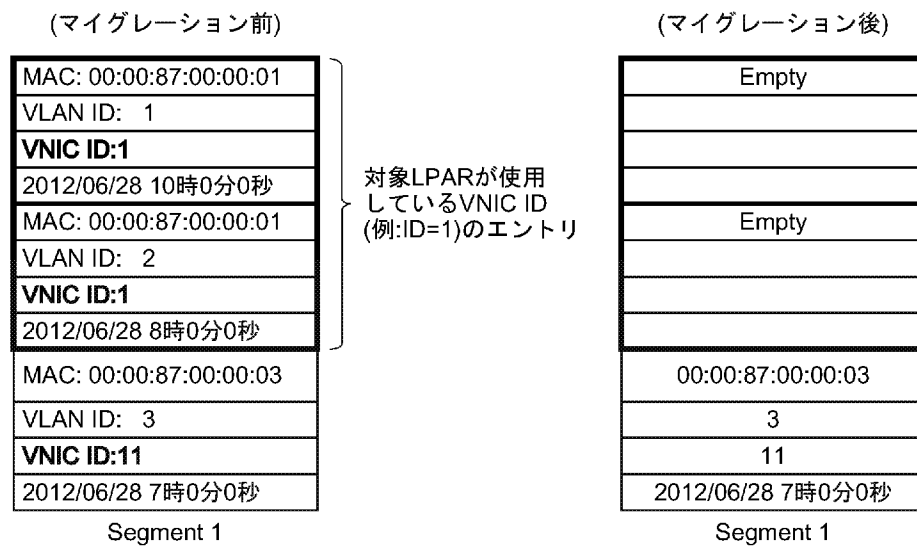
[図11]

Fig. 11



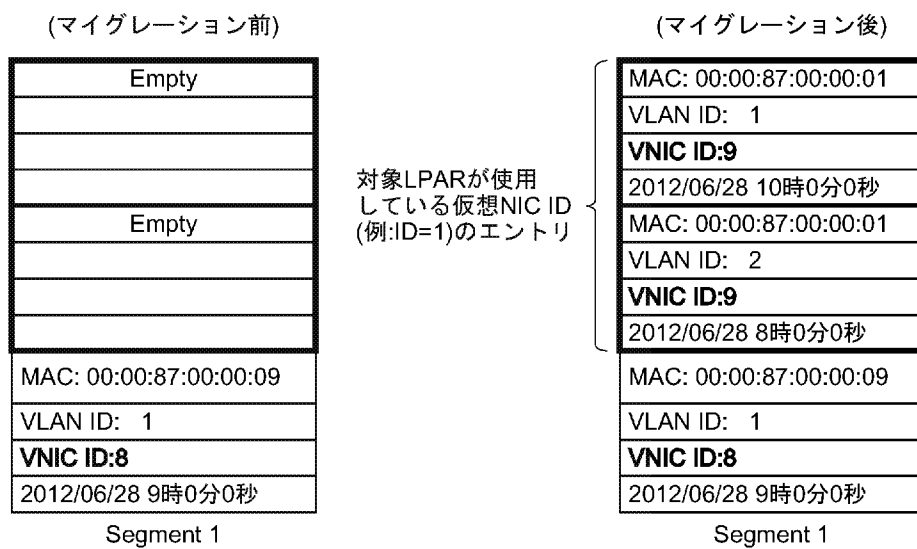
[図12]

Fig. 12



[図13]

Fig. 13



[図14]

Fig. 14

MACアドレス	VLAN	パケット転送先ポート
00:00:87:00:00:01	1(指定なし)	portB

[図15]

Fig. 15

MACアドレス	VLAN	パケット転送先ポート
00:00:87:00:00:01	1(指定なし)	portD

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/071217

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/28 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L12/28

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2012

Kokai Jitsuyo Shinan Koho 1971-2012 Toroku Jitsuyo Shinan Koho 1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2011-198299 A (Fujitsu Ltd.), 06 October 2011 (06.10.2011), paragraphs [0022] to [0102] (Family: none)	1-10

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
11 September, 2012 (11.09.12)

Date of mailing of the international search report
18 September, 2012 (18.09.12)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. H04L12/28 (2006.01) i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. H04L12/28

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 2 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 2 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 2 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2011-198299 A（富士通株式会社）2011. 10. 06, 【0 0 2 2】－【0 1 0 2】（ファミリーなし）	1-10

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

1 1 . 0 9 . 2 0 1 2

国際調査報告の発送日

1 8 . 0 9 . 2 0 1 2

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

岩田 玲彦

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 6

5 X

3 3 6 1