

(51) International Patent Classification:
G06F 15/163 (2006.01)(21) International Application Number:
PCT/MY2011/000079(22) International Filing Date:
6 June 2011 (06.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI 2010005399 16 November 2010 (16.11.2010) MY(71) Applicant (for all designated States except US): **MIMOS BERHAD** [MY/MY]; Technology Park Malaysia, Kuala Lumpur 57000 (MY).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MAT ISA, Mohd Anuar** [MY/MY]; Mimos Berhad, Technology Park Malaysia, Kuala Lumpur 57000 (MY). **MAHMUD, Ramlan** [MY/MY]; Mimos Berhad, Technology Park Malaysia, Kuala Lumpur 57000 (MY). **AB MANAN, Jamalul-Lail** [MY/MY]; Mimos Berhad, Technology Park Malaysia, Kuala Lumpur 57000 (MY).(74) Agent: **NANJAPPAN, Puvaneswari**; Kass International Sdn. Bhd., Suite 8-7-2, Menara Mutiara Bangsar, Jalan

Liku Off Jalan Riong, Bangsar, Kuala Lumpur 59100 (MY).

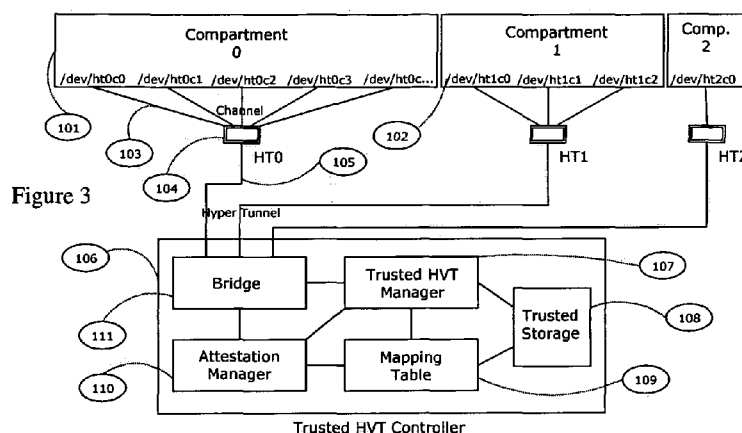
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, 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, RO, RS, RU, 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.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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).

Published:

— with international search report (Art. 21(3))

(54) Title: APPARATUS AND METHOD TO MANAGE INTER-COMMUNICATION BETWEEN COMPARTMENTS USING TRUSTED HYPERVISOR/VISUALIZATION TUNNEL CONTROLLER



(57) Abstract: An apparatus and method that allows communication between different types of compartments using tunnel within or outside the hypervisor/virtualization. These compartments may be created by different virtualization technology. The Virtualization Tunnel realized by a Hyper Tunnel (HT) is a virtual circuit that enables a compartment to communicate with another compartment through manageable inter-communication circuits. The THVT Controller can support numerous compartments running simultaneously and also support diverse type of virtualization platform/technology to have a robust inter-communication without modifying the virtualization platform. For example, compartments from different type of virtualization platforms/technologies such as Virtual Box virtual machine may establish Hyper Tunnel with VMware virtual machine.

**Apparatus and Method to Manage Inter-communication Between
Compartments Using Trusted Hypervisor/Visualization Tunnel
Controller**

5 FIELD OF INVENTION

The present invention relates generally to an apparatus for managing inter-communication between compartments using a trusted hypervisor/visualization tunnel controller and method thereof.

10

BACKGROUND OF INVENTION

All computing systems consist of many individual compartments that in order to work cohesively, must communicate with each other, often in very complex ways and at high speeds. Compartments may include CPUs, memory, monitors, hard disk drives, networks, peripherals, and a host of other devices that combine to form a complete computing system. Very often, these compartments are virtual and the connections done virtually. Currently, there is no general architecture that manages virtual communication efficiently between virtual compartments. There is no standard mechanism to allow different virtual compartments created by different type of virtualization mechanism to have secure and robust inter-communication between them. For example, a compartment created by VirtualBox does not have direct communication with a compartment created by VMware, Qemu, KVM or other type of virtualization. Furthermore, there is no attestation and trusted communication (using TPM) between different compartments in a hypervisor or virtualization.

What is needed is a system that allows manageable inter-communication between compartments.

30

What is also needed is a solution that allows for robust communication between different types of virtualization mechanism/virtualization technology without changing the original virtualization architecture.

- 5 What is further needed is an apparatus that provides attestation and trusted communication between compartments in a hypervisor or virtualization layer.

SUMMARY OF INVENTION

- 10 The present invention aims at providing an apparatus and method for communication between different compartments using tunnel within or outside hypervisor/visualization without changing the original virtualization architecture.

- 15 The present invention aims at providing an apparatus and method that provides attestation and trusted communication between compartments in a hypervisor or virtualization layer via Trusted Computing technology.

- 20 The present invention relates to an apparatus and method that allows communication between different types of compartments using tunnel within or outside the hypervisor/virtualization. These compartments may be created by different virtualization technology. The Virtualization Tunnel realized by a Hyper Tunnel (HT) is a virtual circuit that enables a compartment to communicate with another compartment through manageable inter-communication circuits. However, the administrator or the TPM's owner must
- 25 configure the THVT Controller via the Trusted HVT Manager to allow the communication. Besides that, the Trusted HVT Attestation must perform attestation and secure the communication channel before two or more compartments can establish a Hyper Tunnel. The THVT Controller can support numerous compartments running simultaneously and also support diverse type
- 30 of virtualization platform/technology to have a robust inter-communication without modifying the virtualization platform. For example, compartments from

different type of virtualization platforms/technologies such as VirtualBox virtual machine may establish Hyper Tunnel with VMware virtual machine.

The present invention relates to an apparatus for communication between
5 virtual machine compartments that have been virtualized via more than one
type of virtual machine monitor (VMM) or hypervisor comprising: a plurality of
compartments, with each compartment comprising at least one virtual device; a
controller providing a plurality of connections, each said connection between a
pair of said virtual devices, after attesting to the integrity of said pair of virtual
10 devices based on a predetermined configuration and policy. The controller
comprises: a bridge that connects the virtual devices via communication
channels that have been multiplexed before forming hyper tunnels and then
joining the bridge; a manager module that manages the configuration and
policy of the controller; an encrypted data storage module; a mapping table
15 module consisting of multiple mapping tables including virtual circuit bridge
tables, hyper tunnel tables, and/or channel tables; and an attestation module
for attesting to the integrity of said pair of virtual devices based on a
predetermined configuration and policy before allowing said connection. The
connection between the virtual devices can be any of loop, one to one, one to
20 many, many to one, and many to many.

The present invention also relates to a method for validating a communication
between virtual devices that have been virtualized via more than one type of
virtual machine monitor (VMM) or hypervisor comprising the steps of:

25 a) enabling a controller, said controller providing a plurality of
connections, each said connection between a pair of said virtual
devices;

b) checking if a trusted platform module or virtual trusted platform
module exists;

30 c) if neither said trusted platform module or virtual trusted platform
module exists, disabling said controller; and if either said trusted

platform module or virtual trusted platform module exists, measuring a predefined and stored component; and

d) if said measurement is valid, allowing said communication; and if said measurement is not valid, disabling said controller.

5

The present invention further relates to a method for communication between virtual devices that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor comprising the steps of:

- 10 a) starting a controller, said controller providing a plurality of connections, each said connection between a pair of said virtual devices;
- b) configuring said controller via a manager module, including configuring a boot loader, said boot loader pointing a trusted grub to conduct an integrity measurement on the said controller and related
- 15 objects;
- c) measuring specific program configuration of said controller;
- d) sealing and storing said measurement of c);
- e) rebooting the controller to apply the new measurement;
- f) repeating steps b) to e) as required;
- 20 g) unsealing controller configuration;
- h) loading a mapping table from storage;
- i) attesting to integrity of hyper tunnels;
- j) attesting to integrity of channels;
- k) bridging virtual circuit if attesting steps i) and j) are positive;
- 25 l) mounting virtual device; and
- m) send data through mounted virtual device.

These and other objects of the present invention will become more readily apparent from the detailed description given hereinafter. However, it should be

30 understood that the detailed description and specific examples, while indicating the preferred embodiments of the invention, are given by way of illustration

only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

5 BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows a diagram of a first embodiment of the present invention.

Figure 2 shows a diagram of a second embodiment of the present invention.

10

Figure 3 shows a diagram of system architecture in an embodiment of the present invention.

Figure 4 shows a diagram of relationship and communication between
15 compartments in an embodiment of the present invention.

Figure 5 shows a flowchart of a process to validate and verify the THVT component in an embodiment of the present invention.

20 Figure 6 shows a flowchart of a process for checking and starting the hyper tunnel and the channel in an embodiment of the present invention.

DETAILED DESCRIPTION OF INVENTION

25 It should be noted that the following detailed description is directed to an apparatus and method for communication between different compartments using tunnel within or outside hypervisor/visualization without changing the original virtualization architecture and is not limited to any particular size or configuration but in fact a multitude of sizes and configurations within the
30 general scope of the following description.

The present invention can be implemented in two modes: mode 1 is illustrated in Figure 1 and mode 2 in Figure 2.

Figure 1 shows a diagram of mode 1 in a first embodiment of the present invention. In this mode, the Trusted Hypervisor/Virtualization Tunnel (THVT) Controller may reside within the hypervisor or virtualization. This situation requires a little modification on the hypervisor to support this implementation.

Figure 2 shows a diagram of mode 2 in a second embodiment of the present invention. In this mode, the Trusted Hypervisor/Virtualization Tunnel (THVT) Controller may reside outside the hypervisor or virtualization (e.g. reside in the Operating System, Boot loader, Main Board, BIOS and etc.). This situation requires little or no modification on the hypervisor to support this implementation.

Figure 3 shows a diagram of system architecture in an embodiment of the present invention. This invention uses either a trusted platform module (TPM) or virtual trusted platform module (vTPM) as the root of trust for integrity measurement and other TPM functionality to preserve a chain of trust from a lower layer until a top layer. The compartments or virtual machines (101) are in an isolated area which have been virtualized via a virtual machine monitor (VMM) or hypervisor. Virtual devices (102) within compartment represent a communication channel (103) with other channels. All input/output data will go through these virtual devices. These channels are multiplexed (104) before going through a Hyper Tunnel (HT). The Hyper Tunnel (105) is a multiplexed communication channel with higher data transfer rate so that it can establish secure and trusted communication with the Bridge (111).

The Trusted Hypervisor/Virtualization Tunnel (THVT) Controller (106) and its components could be either physical hardware or virtualized using software. Administrator or the TPM's owner must configure the THVT Controller via the

THVT Manager Module (107). The Trusted Storage Module (108) protects the entire data via a TPM_Seal (e.g. THVT Controller configuration and policies, mapping tables and etc). The Mapping Table (109) has multiple mapping tables such as virtual circuit bridge table, hyper tunnel table, channel table and etc. An Attestation Manager Module (110) performs attestation on platform, compartment (vm machine), THVT Controller(s), hyper tunnel, channel and etc. After the Attestation Manager (110) completes its task, the Bridge (111) will connect the virtual circuits based on configuration and policies.

Figure 4 shows a diagram of relationship and communication between compartments in an embodiment of the present invention, including the virtual relation and communication between compartments via the mapping table. The Hyper Tunnel is represented as tunnel (201) with its id (e.g. HT0, HT1, etc.). The channel is represented as a channel (202) with its id (e.g. C0, C1, etc.). The Relation is represented as a relation (203) with type of relation such as one to one, one to many and etc. The cross links between channel, tunnel and relation provide virtual communication between virtual machines such as (HT0C0, HT1C1) (204, 205).

Figure 5 shows a flowchart of a process to validate and verify the THVT component in the computer system in an embodiment of the present invention. As we know that the trusted boot loader or trusted GRUB will perform integrity measurement to the THVT components and related objects. It then stores (TPM_Extend) this measurement within the TPM in the Program Configuration Register (PCR) array, in locations PCR0 until PCR15 or above. Every time a machine is booting, this process will capture the stored integrity measurement and compares it with the actual measurement. This process will ensure that all THVT components or objects run at the trusted state and thus, maintaining the chain of trust. The process to start the THVT Controller by first verifying the existence of TPM/vTPM in the computer system is as follows: Enable THVT Controller (301) will start the THVT components on the system and

consequently check for the existence of the TPM hardware (302) before booting the THVT Controller. If the hardware TPM does not exist in the computer system, an alternative to TPM hardware, virtual TPM (vTPM) will be checked (303) before booting the THVT Controller. The THVT Controller (304) will be disabled if TPM/vTPM does not exist in the systems. If the THVT Controller is running for the first time, this controller must be configured (or apply for trusted measure) from the trusted boot loader or trusted sequences agents to measure all related objects to this invention. Furthermore, the client must create an administrator account and configure the THVT Controller. The THVT Components is then measured (305), which takes the measurement of the THVT component and compares it with actual measurements (306).

Figure 6 shows a flowchart of a process for checking and starting the hyper tunnel and the channel in an embodiment of the present invention. If it is the first time, the THVT Controller is started (401), and an administrator account is created and the THVT Controller configured (402) via the THVT Manager. The configuration process also includes the process to configure trusted boot loader (pointing trusted grub to do integrity measurement on the THVT Controller and related objects). Furthermore, it then measures the THVT Controller (403) via TPM_ EXTEND to specific PCR and stores this integrity measurement in the trusted storage (404) (TPM_SEAL or TSS_BIND) for later use. The machine must be rebooted in order to apply the new integrity measurement. After the client has completed performing the administration configuration (after reboot), the client needs to re-logon into the system. After the client successfully logs on to this system, the client (administrator) may choose to reconfigure this service (402) repeatedly as required, or continue to perform unseal the THVT Controller configuration (405). The Load Mapping Table (406) will load (TPM_UNSEAL or TPM_UNBIND) original data from the trusted storage. The THVT Attestation must be performed attestation on hyper tunnels (407) and channels (408) to secure the communication channel before two or more compartments could establish the hyper tunnels and channels.

After the Attestation Manager completes its task, the Bridge will connect the virtual circuits (409) based on configuration and policies. Within the virtual machine, the operating system or boot loader could mount the virtual device (410). Finally, the client may start sending and receiving data (411) through
5 mounted virtual devices.

While several particularly preferred embodiments of the present invention have been described and illustrated, it should now be apparent to those skilled in the art that various changes and modifications can be made without departing from
10 the spirit and scope of the invention. Accordingly, the following claims are intended to embrace such changes, modifications, and areas of application that are within the spirit and scope of this invention.

CLAIMS

1. An apparatus for communication between virtual machine compartments that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor comprising:
- 5 a plurality of compartments (101), each said compartment (101) comprising at least one virtual device (102);
- a controller (106) providing a plurality of connections, each said connection between a pair of said virtual devices (102), after attesting to the integrity of said pair of virtual devices (102) based on a predetermined configuration and policy.
- 10
2. An apparatus for communication between virtual machine compartments that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor according to claim 1 wherein the said controller (106) comprises:
- 15 a bridge (111) that connects the said virtual devices (102) via communication channels (103) that have been multiplexed before forming hyper tunnels (105) and joining the said bridge (111);
- 20 a manager module (107) that manages the configuration and policy of the said controller (106);
- an encrypted data storage module (108);
- a mapping table module (109) consisting of multiple mapping tables; and
- 25 an attestation module (110) for attesting to the integrity of said pair of virtual devices (102) based on a predetermined configuration and policy before allowing said connection.
3. An apparatus for communication between virtual machine compartments that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor according to claim 2
- 30

wherein said multiple mapping tables include virtual circuit bridge tables, hyper tunnel tables, and/or channel tables.

- 5 4. An apparatus for communication between virtual machine compartments that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor according to any of the preceding claims wherein the said connection between virtual devices (102) can be any of loop, one to one, one to many, many to one, and many to many.
- 10
5. A method for validating a communication between virtual devices (102) that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor comprising the steps of:
- 15 a) enabling a controller (106), said controller providing a plurality of connections, each said connection between a pair of said virtual devices (102);
- b) checking if a trusted platform module (302) or virtual trusted platform module (303) exists;
- 20 c) if neither said trusted platform module (302) or virtual trusted platform module (303) exists, disabling said controller (106); and if either said trusted platform module (302) or virtual trusted platform module (303) exists, measuring a predefined and stored component (305); and
- 25 d) if said measurement (305) is valid, allowing said communication; and if said measurement (305) is not valid, disabling said controller (106).
- 30 6. A method for communication between virtual devices (102) that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor comprising the steps of:

- 5 a) starting (401) a controller (106), said controller providing a plurality of connections, each said connection between a pair of said virtual devices (102);
- b) configuring (402) said controller (106) via a manager module (107);
- c) measuring (403) specific program configuration of said controller;
- d) sealing and storing (404) said measurement of c);
- e) rebooting the controller to apply the new measurement;
- f) repeating steps b) to e) as required;
- 10 g) unsealing (405) controller configuration;
- h) loading (406) a mapping table from storage;
- i) attesting (407) to integrity of hyper tunnels;
- j) attesting (408) to integrity of channels;
- k) bridging (409) virtual circuit if attesting steps i) and j) are positive;
- 15 l) mounting (410) virtual device; and
- m) send and receive (411) data through mounted virtual device.

- 20 7. A method for communication between virtual devices (102) that have been virtualized via more than one type of virtual machine monitor (VMM) or hypervisor according to claim 6 wherein said configuring (402) further includes configuring a boot loader, said boot loader pointing a trusted boot loader to conduct an integrity measurement on the said controller and related objects.

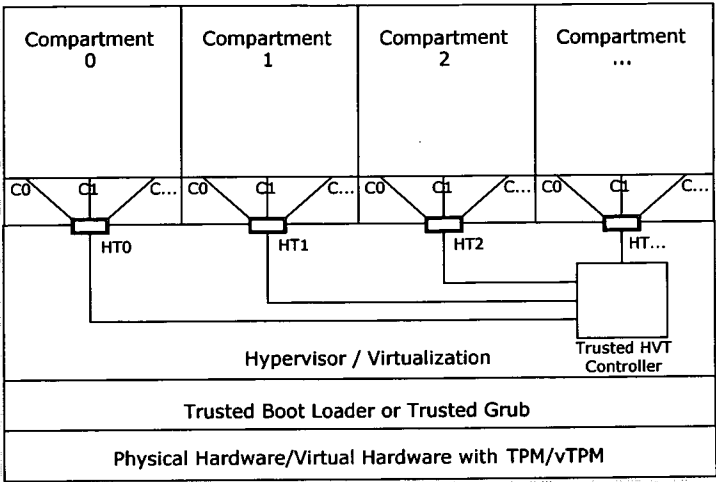


Figure 1

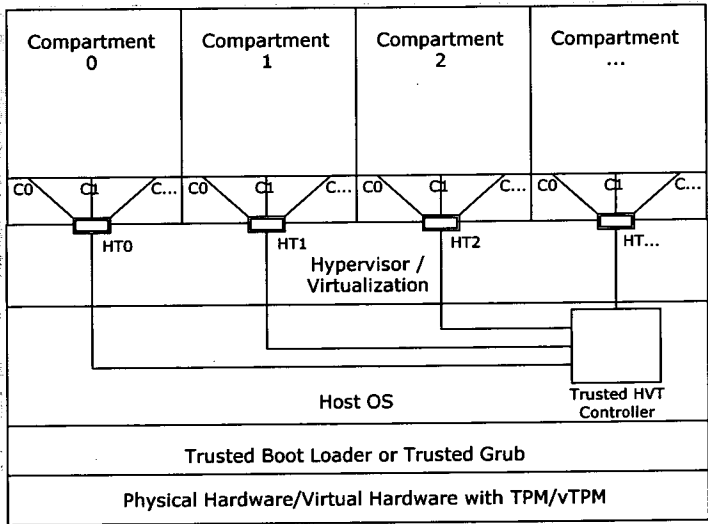


Figure 2

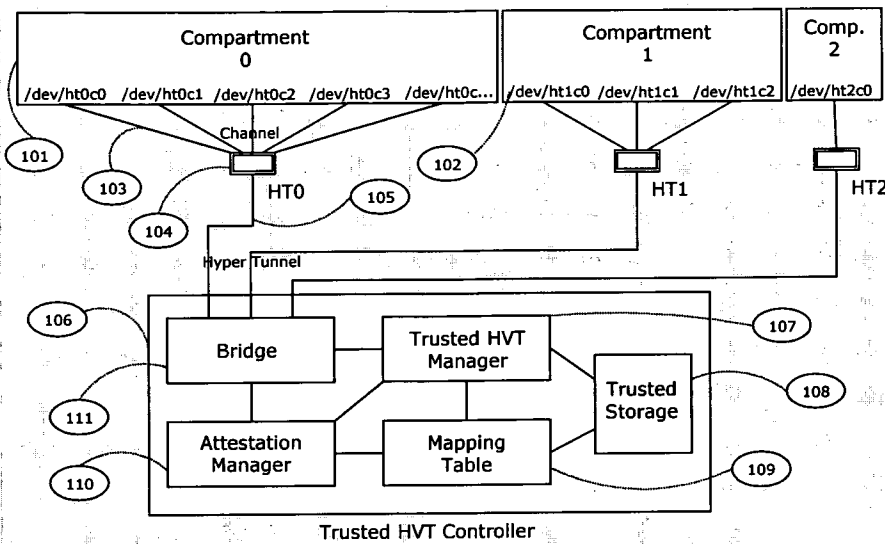


Figure 3

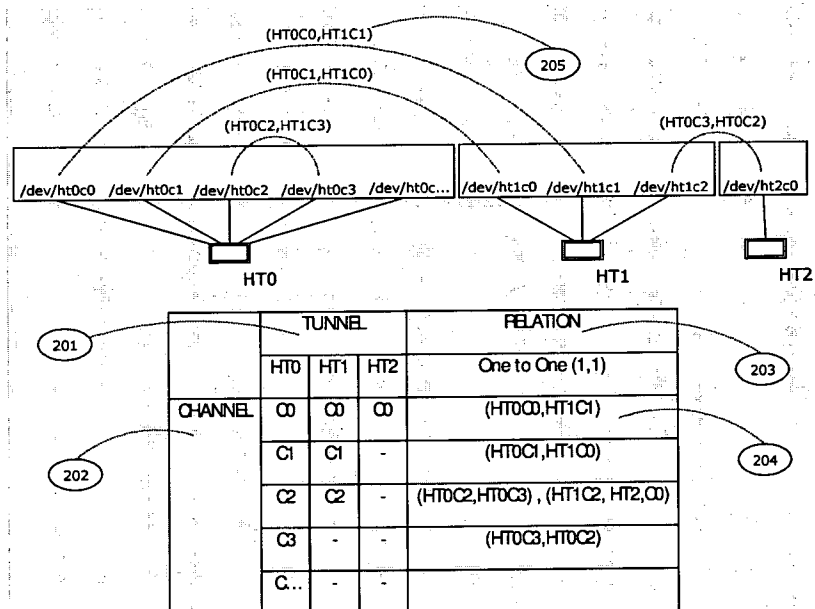


Figure 4

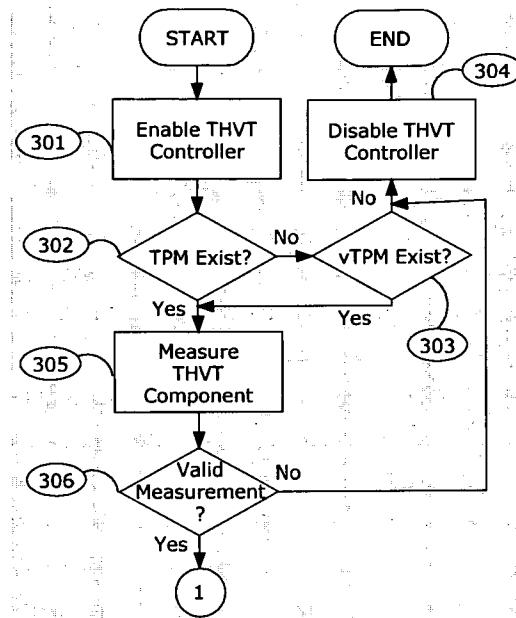


Figure 5

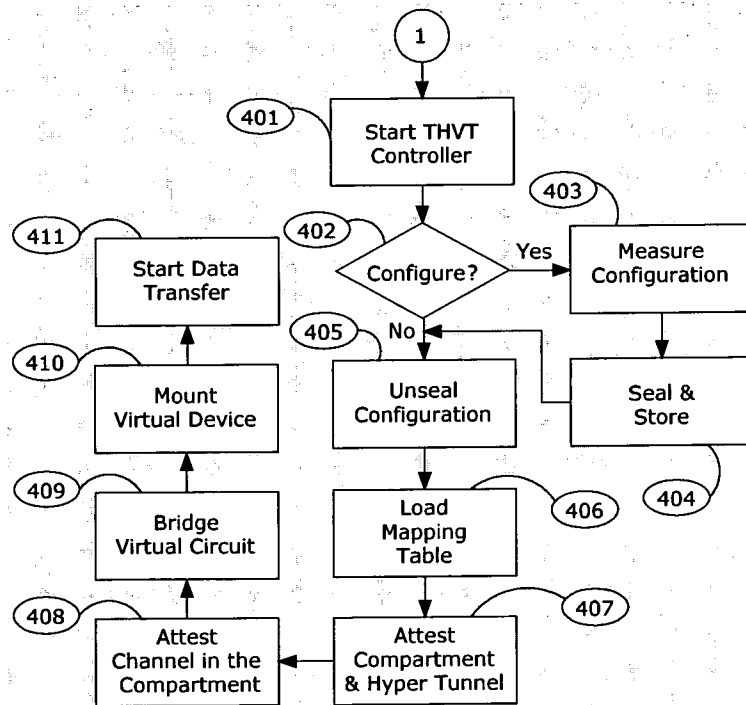


Figure 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/MY2011/000079

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G06F 15/163 (2006.01)

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)

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

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

File EPODOC & WPI with Keywords (VIRTUAL MACHINE MONITOR, COMPARTMENT, PARTITION) and like terms

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2005/0246521 A1 (BADE et al.) 3 November 2005	
A	US 2009/0007104 A1 (ZIMMER et al.) 1 January 2009	

☐

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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
21 November 2011

Date of mailing of the international search report
25/11/2011

Name and mailing address of the ISA/AU
AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustalia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer
Vivek Joshi
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 2 6222 3663

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/MY2011/000079

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member
US	2005246521	NONE
US	2009007104	NONE
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.		
END OF ANNEX		