Heartbeat and Drbd the application in the large capacity OLT

WangHe SunXiaoyan ZhangZenghua LiChunlong WangXiangqun

The State Grid Electric Power Research Institute，Nanjing Jiangsu 211100;

Abstract： With Drbd OLT system for large-capacity high-availability performance requirements , to Heartbeat as the research object , the system with hot standby high availability also has the function of the synchronization file. This paper discussed the basic principles and feasibility of Heartbeat and Drbd application , to achieve the thermal systems used in large-capacity OLT , improved switching speed , and increase its stability . Actual tests show Heartbeat and Drbd can be well applied to the large capacity management OLT , rapid and stable system to meet the high availability of applications. And reduce costs, based on free software under the open source Linux-HA project makes it has a higher market competitiveness

Keywords： Linux Heartbeat High Availability Synchronous Drbd

1. Introduction

The emergence of a high availability cluster[1-2] is as far as possible in order to make the whole of the cluster service is available, to reducing the error of computer hardware and software. Large capacity OLT equipment master a master plate adopts a way [3], when the master control board motherboard outage or failure, the standby board in the main control board adopt the way of a Ping packet monitoring to the main control board motherboard failure, spare board start programs (mainly protocol module related procedures), realizes the switch, to improve the availability of the program.
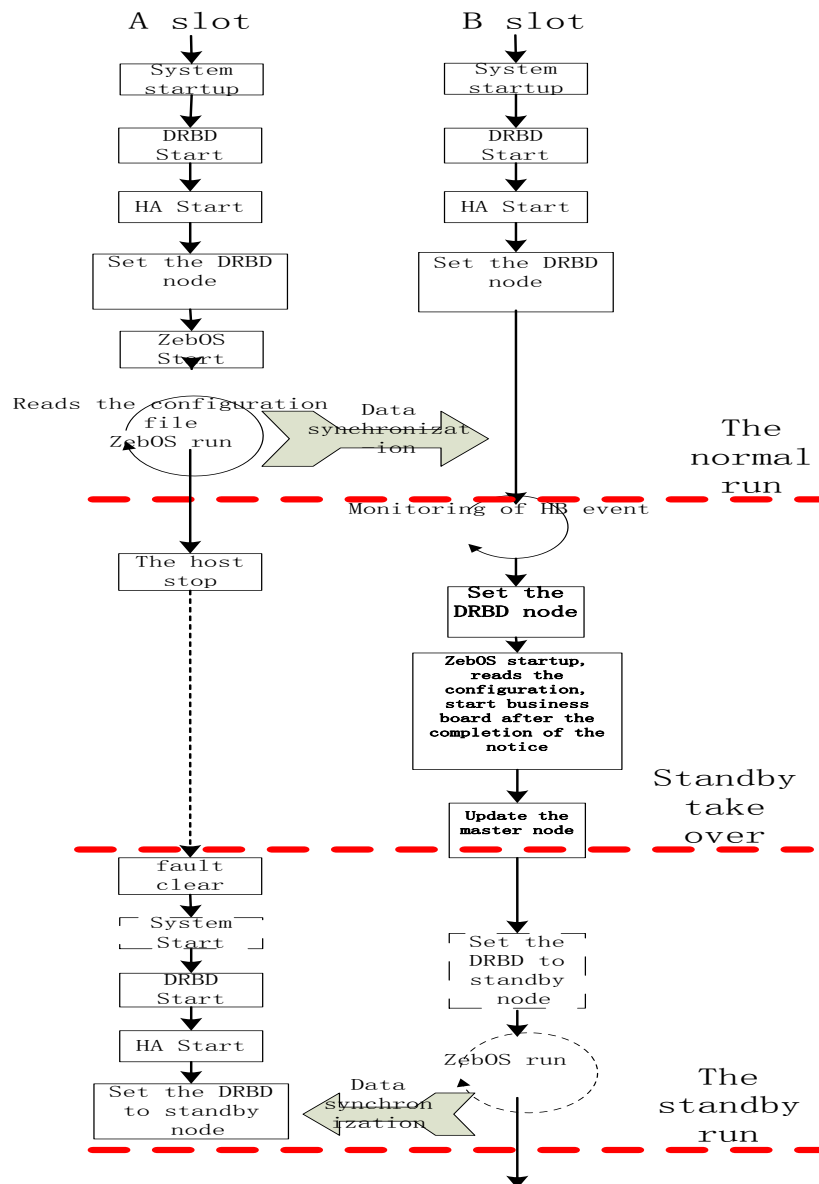
2. Large capacity OLT hot standby system architecture

2.1. Hot backup and synchronization system architecture

Heartbeat hot standby system architecture of in D8000 large capacity OLT using Linux-HA program based on Heartbeat set up a hot standby platform , its work mode from the way (Active / Standby), which is the main control board, the standby board in monitoring the ready state. When the Heartbeat detects that the main control board is down, the standby board over the main control board, the main board is recovered will be prepared plate rather than switching to the motherboard, improve efficiency. The hot standby system using two main control board(D8000 are equipped with double network card) as the main board, Eth0 card access external network of two main control board, for external access; Two pieces of main control board Eth1 card direct interconnection in D8000, internal IP, for use in the D8000 heart; the main control board and the standby board is connected directly as the heart line to connect between the two board. It is in order to avoid unnecessary handover failure, so use a separate physical path as a communication path to

understand the operation of each other. The key of this scheme for fault tolerance function is switched on the motherboard is down, the client is the master server is transparent, i.e. the motherboard switching of the working end doesn't seem to change, all the applications based on the motherboard are normal. Drbd is a piece of equipment, used for high availability (HA). When writing data to the main control board, data also will be sent to the standby board. Recorded in the same in the standby board. The master node and the node data can ensure the real-time synchronization. When the master node of system failure, preparing the node will retain a copy of the same data, can continue to use. The system flow as shown in Figure 1.

```
      A slot                    B slot

  ┌─────────────┐          ┌─────────────┐
  │   System    │          │   System    │
  │   startup   │          │   startup   │
  └─────────────┘          └─────────────┘
         │                        │
  ┌─────────────┐          ┌─────────────┐
  │    DRBD     │          │    DRBD     │
  │    Start    │          │    Start    │
  └─────────────┘          └─────────────┘
         │                        │
  ┌─────────────┐          ┌─────────────┐
  │  HA Start   │          │  HA Start   │
  └─────────────┘          └─────────────┘
         │                        │
  ┌─────────────┐          ┌─────────────┐
  │ Set the DRBD│          │ Set the DRBD│
  │    node     │          │    node     │
  └─────────────┘          └─────────────┘
         │
  ┌─────────────┐
  │    ZebOS    │
  │    Start    │
  └─────────────┘

Reads the configuration
         file              Data
      ZebOS run         synchronization      The
                                            normal
- - - - - - - - - - - - - - - - - - - - -    run
              Monitoring of HB event

  ┌─────────────┐          ┌─────────────┐
  │  The host   │          │   Set the   │
  │    stop     │          │  DRBD node  │
  └─────────────┘          └─────────────┘
                           ┌─────────────┐
                           │    ZebOS     │
                           │  startup,    │
                           │  reads the   │
                           │configuration,│
                           │start business│
                           │board after the│
                           │completion of  │
                           │  the notice   │
                           └─────────────┘
                           ┌─────────────┐
                           │  Update the  │    Standby
                           │ master node  │     take
                           └─────────────┘     over
- - - - - - - - - - - - - - - - - - - - -
  ┌─────────────┐          ┌─────────────┐
  │   fault     │          │  Set the     │
  │   clear     │          │  DRBD to      │
  └─────────────┘          │  standby      │
  ┌─────────────┐          │  node         │
  │   System    │          └─────────────┘
  │   Start     │
  └─────────────┘
  ┌─────────────┐             ZebOS run
  │    DRBD     │
  │    Start    │
  └─────────────┘
  ┌─────────────┐
  │  HA Start   │
  └─────────────┘          Data
  ┌─────────────┐       synchronization      The
  │ Set the DRBD│                           standby
  │ to standby  │                             run
  │    node     │
  └─────────────┘
- - - - - - - - - - - - - - - - - - - - -
```

The heartbeat configuration in the system

Heartbeat uses a virtual IP address mapping techniques to realize this function, the Heartbeat configuration file is the key of the ha, cf and haresources of configuration. Ha. Cf keepalive for 2 seconds, too small can aggravate the burden of the system, accurate, is too big to test in the OLT deadtime for 8 seconds, so make the switch in the packet loss is severe, small error rate increase.

2.2. Communication of data synchronization and installation

Data synchronization software using DRBR (Distributed Replicated Block Device), Drbd is a piece of equipment, can be used for high availability (HA).

When writing data to the local file system, data will be sent to another host on the network. Recorded in the same on a file system. The master node and the node data can ensure the real-time synchronization. When the master node of system failure, preparing the node will retain a copy of the same data, can continue to use.

This scheme adopts the initialization process of ZebOS restart when switching to ZebOS procedure, ensure the accuracy of data read program. The disadvantage is the way is long switching time, the advantage is not re - initialization parameters in program execution, read the file reliable, existing

programs on ZebOS without modification. D8000 need to synchronize the data into two classes:

One is the configuration file is written to the permanent preservation of Flash, users need to input the Save command to save in the Flash, but also into the spare board through the data synchronization software Flash. The data of main control board in the power-down data exists, write Flash slow.

Another type of data configuration data during the operation, the data is generated by the user through the configuration, without the need for permanent preservation, this kind of configuration files temporarily saved to ramdisk. Control panel power-down data loss, but the writing speed. The following is a disk data synchronous data flow, ramdisk data synchronization principle is the same. Principle as shown in Figure 2
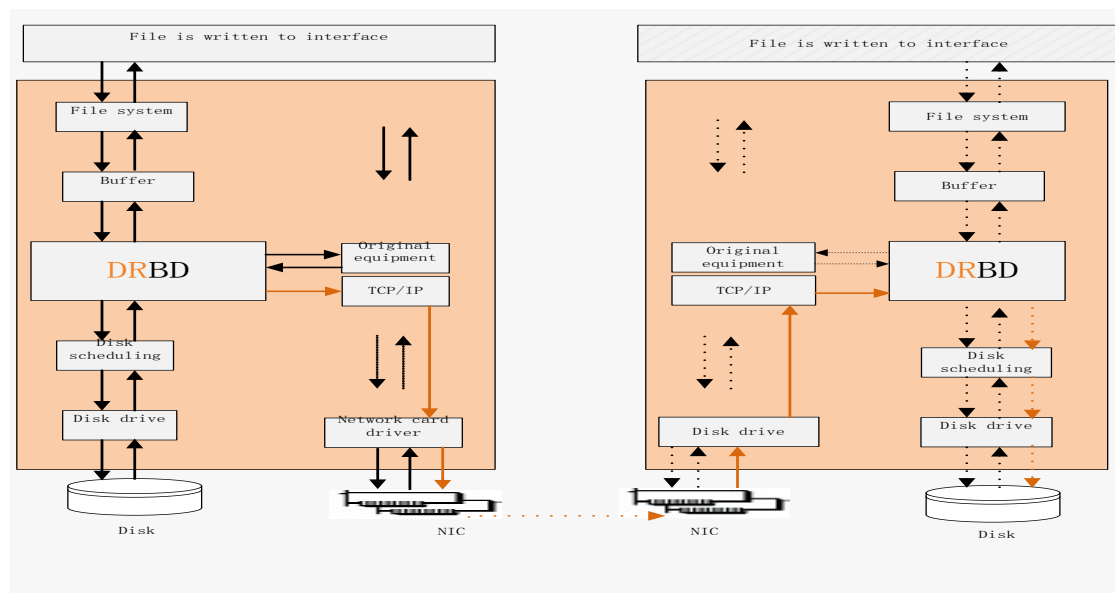
2.2.1 The design requirements of Drbd D8000

In the single node dual mode, system operation is divided into a state.The normal operation of the A slot: the main control board, B control board monitoring slot. A slot for synchronous data stream.

The board took over the slot: A main

control board, main control board B slot node, B slot for synchronizing data source, the B slot by upgrading.

The board operation: A slot of main control board fault recovery, the A slot of main control board to board, B slot for synchronizing data source

2.2.2 The target file Drbd cross compiled in D8000

The generated files are compiled

Drbd.o: drive module

Drbd: start the Drbd service script.

Drbd.conf: system configuration file.

Drbd setup: the generation of executable file, call the Drbd script.

The main control board in the master1, the interface is as follows:

To perform the same operation on the main control panel on the standby1. Note hostname is set to standby1, the specific configuration can be modified in the drbd.conf file

2.3. Heartbeat combined with Drbd scripts required

Heartbeat needs to monitor and control the Drbd, primary and secondary nodes, nodes can only set to primary mount, the master node as the data source, set heartbeat to complete, must write TransferSwitch scripts

```
#!/bin/sh
unset LC_ALL; export LC_ALL
unset LANGUAGE; export LANGUAGE
prefix=/home/mtd2/bin/ha
exec_prefix=/home/mtd2/bin/ha
. /home/mtd2/bin/ha/etc/ha.d/shellfuncs
case "$1" in
'start')
/home/mtd2/bin/ha/ha_client
```

```
Nari DPN8000 HA send message to ZebOS key! 1
/bin/drbdsetup /dev/nbd/0 primary #升级为主节点
/bin/mount /dev/nbd/0 /home/mtd1 #设置挂载
;;
'pre-start')
;;
'post-start')
;;
'stop')
/bin/umount /home/mtd1 #卸载
/bin/drbdsetup /dev/nbd/0 secondary #设置为备用节点点
/home/mtd2/bin/ha/ha_client
** DPN8000 HA send message to ZebOS key! 0
;;
'pre-stop')
;;
'post-stop')
;;
echo
"Usage:$0{start|pre-start|post-s
```

```
tart|stop|pre-stop|post-stop}"
   ;;
   esac
   exit 0
```

3. Summary

In this paper, the design of practical technology OLT large capacity heat based on heartbeat system, with high availability and improve the real-time data processing software Drbd, the organic combination of the two. An improved initialization methods, to read the file allocation method, existing programs on ZebOS without modification, more stable and reliable. The reliability of the system have great.

4. Acknowledgements

5. References

[1] Luo Huang LI Dafeng. Usability principles of measurability and test method [J]. Microcomputer, 2007,05-3:249-251

[2] Changsheng Xie Qingping Hu, Zhihu Tan a new real-time heart rate monitoring technology engineering and computer science in 2004, 26 (5).

[3] SOBELL G M.L inux command, editor, and shell programming [M]. Yang Mingjun, Wang Fengqin, translation. Beijing: tsinghua university press, 2007