
Subject: Re: A question about Node RAM
Posted by [Tim Small](#) on Sat, 07 Jan 2012 17:02:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 06/01/12 22:59, jjs - mainphrame wrote:

> On Fri, Jan 6, 2012 at 12:34 PM, Tim Small <tim@seoss.co.uk
> <mailto:tim@seoss.co.uk>> wrote:
>
>
> pacemaker+heartbeat
>
>
> Interesting idea, I wonder about the tradeoffs. I tend to keep the
> host node pretty lean and run heartbeat/corosync/pacemaker in the CTs,
> if anywhere.
>

We have a few machines where we put the OpenVZ container backing stores on drbd and use heartbeat+pacemaker (we had some issues with corosync during testing when we initially set things up a few years ago, but it's probably fine now) to manage the OpenVZ containers as cluster resources.

Disk writes are relatively expensive so it's not perfect for all workloads, but it works well overall, and has survived real hardware failures (e.g. motherboard failure) with minimal downtime.

It also allows you to move nodes around easily and should allow you to carry out things like host node kernel updates without bringing down containers (using live migration to other HNs) - although we've not gotten around to testing this.

Our machines are in pairs, but really it'd be better to have them in something like groups of four, so that when a HN fails, the remaining 3 HNs each end up running a third of the evicted containers... This would require corosync instead of heartbeat of course (heartbeat supports 2 nodes only).

Tim.

--

South East Open Source Solutions Limited
Registered in England and Wales with company number 06134732.
Registered Office: 2 Powell Gardens, Redhill, Surrey, RH1 1TQ
VAT number: 900 6633 53 <http://seoss.co.uk/> +44-(0)1273-808309
