
Subject: *CLOSED* softirq stats for a VE
Posted by [Pradeep Padala](#) on Thu, 15 Feb 2007 20:17:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,

Currently, /proc/vestat does not show softirq stats for a particular VE. Is there any way to get those stats ? Please let me know.

Thanks,
Pradeep

Subject: Re: softirq stats for a VE
Posted by [Vasily Tarasov](#) on Fri, 16 Feb 2007 07:41:02 GMT
[View Forum Message](#) <> [Reply to Message](#)

As far as know, there is no such possibility.
/proc/stat in VE just output 0 for soft interrupt ticks.

Vasily

Subject: Re: softirq stats for a VE
Posted by [dev](#) on Mon, 19 Feb 2007 21:07:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

softirqs do system-specific job and it is very hard if possible at all to distinguish to which VE softirq can be charged to. So currently it is reported as 0.

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [Pradeep Padala](#) on Mon, 05 Mar 2007 15:07:02 GMT
[View Forum Message](#) <> [Reply to Message](#)

I know that softirqs are difficult to account for. I am running a network heavy process (apache queried by httpperf) and it produces enormous amount of CPU consumption is softirqs. This obviously has a lot of implications as I cannot control the VE consumption directly.

Do you think the following patch for more accurate process accounting might help ?

<http://www.kip.uni-heidelberg.de/ti/HLT/software/software.html#kernel>

Please let me know. I am happy to make it work with OpenVZ, if it has any promise.

Pradeep

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [dev](#) on Mon, 05 Mar 2007 15:50:04 GMT
[View Forum Message](#) <> [Reply to Message](#)

No, this patch doesn't help. It's just make accounting more accurate (i.e. improves precision of accounting).

Look, the problem is the following. Imagine that someone does a network flood at your host.

It takes a significant CPU time to determine to which VE (if any at all) these packets are going to. And this time is not accounted to the VE. That's the problem. The same for IRQ. IRQ doesn't know anything about whom it is reading packets for. So IRQ time can not be accounted to the VE.

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [Pradeep Padala](#) on Mon, 05 Mar 2007 19:47:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

I understand that when VEs receive data, it's difficult to account for the network processing. But, if the VEs are sending data, wouldn't the network stack know how to account for the CPU used for processing that ?

Pradeep

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [dev](#) on Tue, 06 Mar 2007 08:06:42 GMT
[View Forum Message](#) <> [Reply to Message](#)

The time spend by the VE by sending data is accounted to the VE.
However, when the packets gets into the host system for the routing or switching, it is now processed in VE0 context and this time is accounted for the router, which looks reasonable enough.

Looks like you want to account the whole time spent on packet handling until it gets out of ethX device to the corresponding VE?

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [Pradeep Padala](#) on Tue, 06 Mar 2007 14:17:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

Right. To give you the context, we have been working on developing adaptive control in virtual environments. A paper on this is being published at EuroSys (<http://www.eecs.umich.edu/~ppadala/pubs/eurosys.pdf>). We did that for CPU and now want to

extend it to the network. We used Xen in that work, but are switching to OpenVZ, as we think it has much lower overhead.

When a network-processing-heavy application runs in a VE, it consumes both network and CPU resources. We only want it to be accounted properly. Perhaps I need some thing like lazy receiver processing (<http://www.cs.rice.edu/CS/Systems/LRP/>).

Any ideas are welcome.

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [dev](#) on Tue, 06 Mar 2007 15:06:22 GMT
[View Forum Message](#) <> [Reply to Message](#)

well... some ideas:

1. time should be accounted in cycles, as jiffies are too much inaccurate for such kind of accounting.
 2. From net to VE path: it is possible to mark skb with cycles timestamp on it's arrival. When skb finally arrives to `venet_xmit`, `delta_time = current_cycles() - skb->arrival_cycles` can be accounted to the VE.
 3. on VE to net path: skb can be marked with VEID in `venet_xmit()` when it is moved from VE to VE0 and with a cycles timestamp.
Before doing `hard_start_xmit` the same delta time as above can be accounted to appropriate VE.
 4. Both (2) and (3) work fine if skb is not queued anywhere for delayed sending (e.g. traffic shaping). On delayed sending time accounting can be done partially: time used before queueing and time used after queueing.
-

Subject: Re: *CLOSED* softirq stats for a VE
Posted by [Pradeep Padala](#) on Tue, 06 Mar 2007 15:53:40 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks for the ideas. I might hack the code, when I get time. Will send a patch, if I do.
