Using Tracing to Analyze Hard Disk Performance

Houssem Daoud
December 10, 2015

École Polytechnique de Montréal
Laboratoire DORSAL
Agenda

Introduction and objectives

Implementation

- Block Layer Structure
- Relevant Tracepoints
- State History Tree

Developed Views

Use Cases – Demo

- Writeback Algorithm
- Lttng I/O Behavior
- Flush requests
Introduction

The performance of block devices has a big impact on overall system performance.

Many optimization techniques have been developed.

Scheduling
- Request sorting
- Request merging

Caching
- Page cache
- Writeback
Introduction

The Block Layer is very complex

It is difficult to detect disk I/O problems

The need to develop an efficient tool to analyze block layer behavior
Introduction and Objectives

Use tracing to Analyse Hard Disk Performance

- Insert Tracepoints in the block device layer
- Create views to visualize important metrics
Implementation

Block Layer Structure

Figure 1: Linux block layer structure
Implementation

Relevant tracepoints

- lttng_statedump_block_device
- block_rq_insert
- block_rq_merge
- block_rq_issue
- block_rq_complete
Implementation

State History Tree

Disks
- sda
  - Elevator queue
    - position
      - Request
        - Size
        - Flags
  - Dispatch queue
- sdb
Developed Views

Waiting Queues

Elevator Queue

Dispatch Queue

Write

Read
Developed Views

Waiting Queues

Merging
Developed Views

Waiting Queues

Queue Length

![Graph showing dispatch queue length over time with peaks and troughs indicating request activity](Image)
Developed Views

Request Latencies

Latency Average
Developed Views

Request Latencies

Latency Distribution

![Latency Distribution chart]

- Latency range [166μs, 229μs]: 0 requests
- Latency range [229μs, 292μs]: 0.87 requests
- Latency range [292μs, 355μs]: 17.04 requests
- Latency range [355μs, 419μs]: 59.91 requests
- Latency range [419μs, 482μs]: 20.56 requests
- Latency range [482μs, 545μs]: 0.77 requests
- Latency range [545μs, 609μs]: 0.41 requests
Developed Views

Disk throughput

![Disk I/O View](image_url)

- Time
- MB/s
# Developed Views

## I/O Activity of Processes

<table>
<thead>
<tr>
<th>PID</th>
<th>Process</th>
<th>Read (MB)</th>
<th>Write (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1724</td>
<td>Xorg</td>
<td>0.0</td>
<td>0.051</td>
</tr>
<tr>
<td>2595</td>
<td>cinnamon</td>
<td>0.0</td>
<td>0.004</td>
</tr>
<tr>
<td>8315</td>
<td>lttnget</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>6271</td>
<td>/usr/bin/termin</td>
<td>0.0</td>
<td>0.022</td>
</tr>
<tr>
<td>2612</td>
<td>nemo</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1474</td>
<td>lqbalance</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8212</td>
<td>/usr/bin/termin</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2029</td>
<td>collect</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3194</td>
<td>chrome</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3014</td>
<td>chrome</td>
<td>0.0</td>
<td>0.007</td>
</tr>
<tr>
<td>6761</td>
<td>lttnget-consumerd</td>
<td>0.0</td>
<td>0.206</td>
</tr>
<tr>
<td>8247</td>
<td>lttnget-simple</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8316</td>
<td>hdparm.sh</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8317</td>
<td>sudo</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>943</td>
<td>rsyslogd</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2820</td>
<td>accounts-daemon</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8318</td>
<td>hdparm</td>
<td>1548.0</td>
<td>0.0</td>
</tr>
<tr>
<td>5328</td>
<td>java</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2446</td>
<td>rtki-daemon</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3206</td>
<td>chrome</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2615</td>
<td>thunderbird</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3619</td>
<td>chrome</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3149</td>
<td>chrome</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4197</td>
<td>okular</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4153</td>
<td>oed</td>
<td>0.0</td>
<td>0.001</td>
</tr>
</tbody>
</table>

![Graph showing I/O activity of processes](image)
DEMO

• Flush requests
• Writeback Algorithm
• Lttng I/O Behavior
Thank You!