



Bridging perf and ftrace: Recording perf in ftrace buffers and perf data to trace.dat conversion

Tanushree Shah <tshah@linux.ibm.com>, IBM
Madhavan Srinivasan <maddy@linux.ibm.com>, IBM
Steven Rostedt <rostedt@goodmis.org>

Linux Plumbers Conference 2026 Tracing MC



LINUX
PLUMBERS
CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026

Motivation

Why convert perf.data into trace.dat ?

perf ecosystem

Statistical profiling + counters
Captures tracepoint events

Output: **perf.data**

trace-cmd ecosystem

Event timeline visualization
Graphical support in KernelShark
Per-CPU event streams

Output: **trace.dat**

Both ecosystems capture the same tracepoint events - but store them in incompatible binary formats, forcing users to record twice.

- ✗ Cannot visualize perf data in KernelShark
- ✗ Redundant recordings in constrained environments
- ✗ No cross-toolchain data sharing



Implementation

Core trace.dat export infrastructure

Introduces new infrastructure within perf implementing per-CPU raw event buffer management, ftrace ring buffer page construction, and section writers for the strings, options, and flyrecord sections of the trace.dat v7 format

Reconstruct trace.dat metadata from perf.data

Writes initial format header (magic, version, endian, page size) and section headers using format data embedded in perf.data at record time.

Process tracepoint samples and serialize to trace.dat

Processes PERF_TYPE_TRACEPOINT samples per-CPU, writes Options (CPUCOUNT, TRACECLOCK, metadata offsets) and flyrecord sections with per-CPU ring buffer pages.

Proposed option --to-trace-dat flag in perf data-convert

Extends builtin-data.c with the new option : mutually exclusive with --to-ctf and --to-json, consistent with existing perf convert options

Patch series: <https://lore.kernel.org/linux-perf-users/20260822062729.882005-1-tshah@linux.ibm.com/>



LINUX

PLUMBERS CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026



Usage

```
#perf record -e sched:sched_switch -e sched:sched_wakeup -a sleep 10 ->
```

```
#perf data convert --to-trace-dat=trace.dat ->
```

```
#kernelshark trace.dat
```

```
tanushree:perf$ sudo ./perf record -e sched:sched_switch -e sched:sched_wakeup -a sleep 10
[sudo] password for tanushree:
[ perf record: Woken up 203 times to write data ]
[ perf record: Captured and wrote 55.816 MB perf.data (480464 samples) ]
tanushree:perf$ sudo ./perf data convert --to-trace-dat=trace.dat
[ perf data convert: Converted 'perf.data' into trace.dat format 'trace.dat' ]
[ perf data convert: Converted 480464 events ]
tanushree:perf$ trace-cmd dump --summary trace.dat
```

Tracing meta data in file trace.dat:

[Initial format]

```
7      [Version]
0      [Little endian]
8      [Bytes in a long]
4096   [Page size, bytes]
none   [Compression algorithm]
       [Compression version]
```

[buffer "", "local" clock, 4096 page size, 22 cpus, 28946432 bytes flyrecord data]

[10 options]

[Saved command lines, 0 bytes]

[Kallsyms, 0 bytes]

[ftrace format, 0 events]

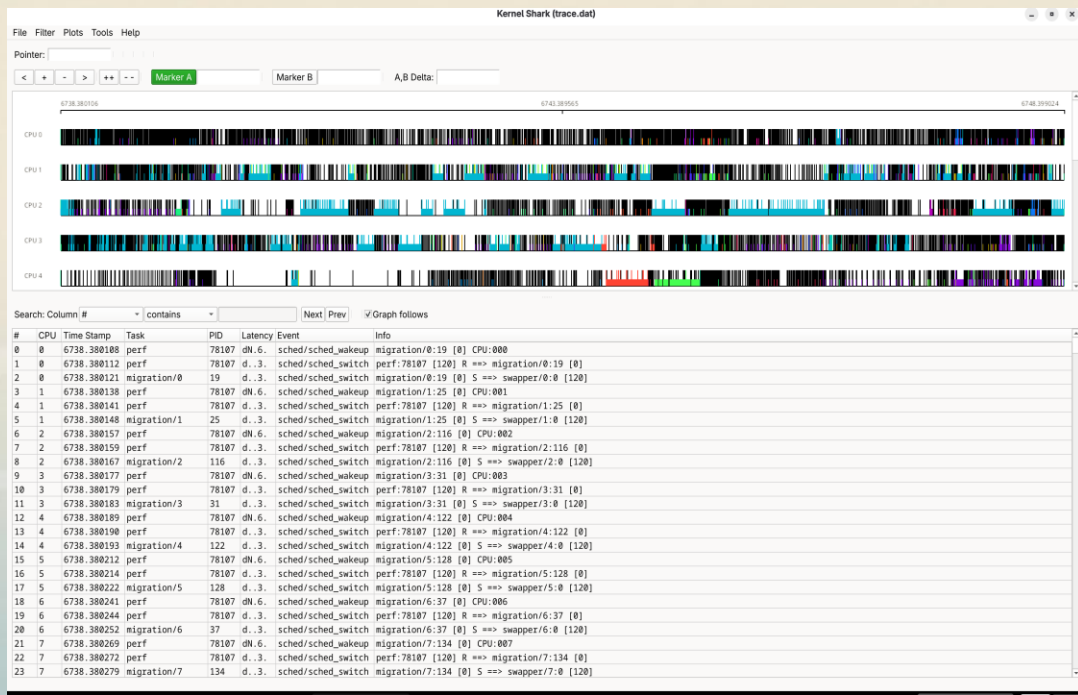
[Header page, 205 bytes]

[Header event, 205 bytes]

[Events format, 1 systems]

[9 sections]

tanushree:perf\$



LINUX

PLUMBERS CONFERENCE PRAGUE, CZECHIA / OCT. 5-7, 2026

Hurdles to work through for future work

No event format for non-tracepoints

ftrace readers expect a field-layout "format" string per event type; Hardware PMU and software counter events have none today.

Compression compatibility

Two paths: link libtracecmd vs. reimplement independently (no dependency but must exactly replicate section layout and version negotiation)

Performance on large-scale traces

Per-CPU buffer management and flyrecord serialization characteristics under high event volumes.



Legal Statement

- This work represents the view of the authors and does not necessarily represent the view of the employers.
- IBM and IBM (Logo) are trademarks or registered trademarks of International Business Machines in United States and/or other countries.
- Linux is a registered trademark of Linus Torvalds.
- Other company, product and service names may be trademarks or service marks of others.

