

Vizualizace trasování procesů v Linuxu

Autor: David Jaromír Šebánek
2025

Vedoucí: RNDr. Jan Kára, Ph.D.
Matematicko-fyzikální fakulta, Univerzita Karlova

Úvod

Operační systémy a procesy, které na nich běží, je často nutné neustále kontrolovat a snažit se odstranit či omezit různé výkonnostní překážky. Jedním ze způsobů analýzy chování systému je skrze **trasování**. Po sbírání trasovacích dat jsou pak pro analýzu často nezbytné **vizualizace** těchto dat. Na Linuxu existuje několik nástrojů jak pro sbírání dat, tak pro jejich vizualizaci, některé jsou i součástí jádra operačního systému. Jednu z těchto vizualizačních aplikací v této práci rozšíříme.

KernelShark

KernelShark je open-source aplikace vyvinuta jako vizualizátor dat z Trace-cmd. Zobrazuje události procesů na časové ose společně se seznamem detailů o událostech.

Aplikace je postavena na GUI frameworku Qt, její zdrojový kód je pak napsán hlavně v jazycích C++ a C. Funkcionality jsou buď součástí kódu aplikace, některé lze dodat skrze rozhraní pro pluginy.

Trasování na Linuxu

Sběr trasovacích dat podporuje Linux široce. Pro tuto práci zajímavými pak byly *tracepointy*, *tracefs*, *Ftrace* a hlavně **Trace-cmd**. Častým nástrojem, ovšem na statické profilování, je program *Perf*. Za zmínku stojí i projekt *LTTng*.

Nasbíraná data jsou pak vizualizována pomocí dalších programů, aby bylo jednodušší je analyzovat člověkem. V této oblasti existuje také několik projektů, například *HPerf*, *TraceShark* či *FlameGraphs*. Práce se ovšem zaměřovala na frontend pro Trace-cmd, program jménem **KernelShark**.

Cíle

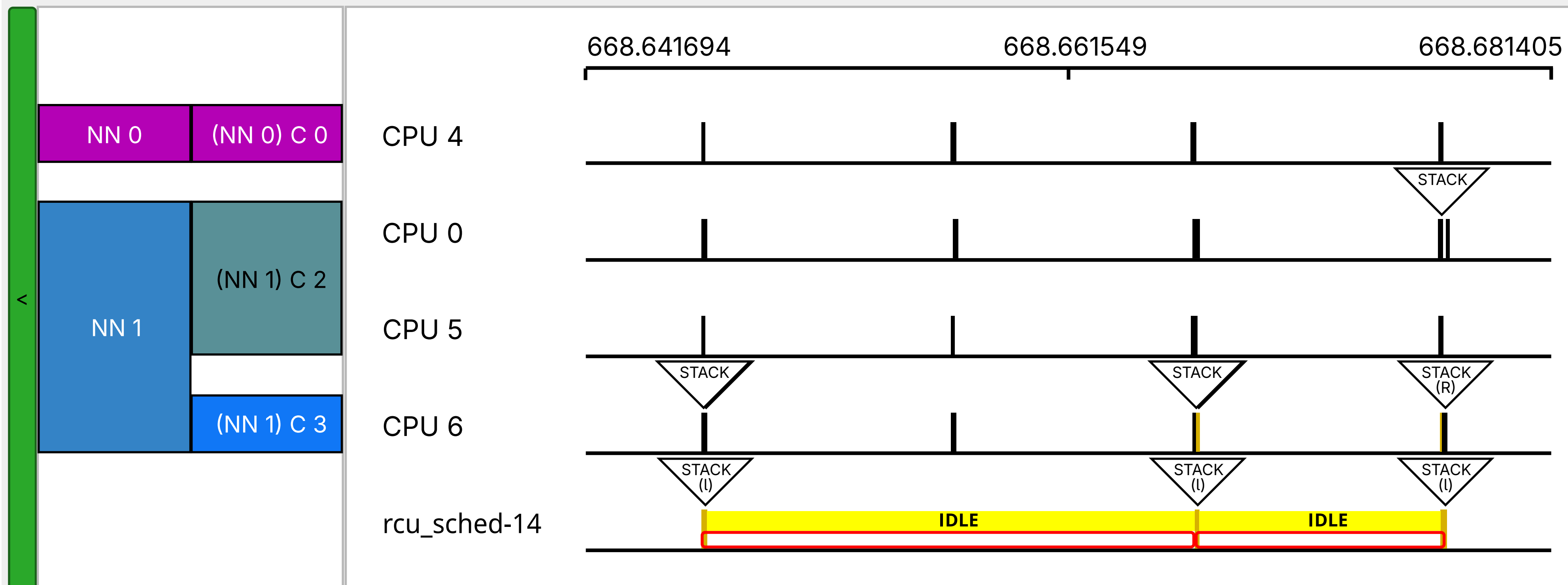
Obečným softwarovým cílem této práce je rozšíření aplikace Kernelshark za účelem zlepšení analýzy systému:

- ♦ přidání GUI prvků pro příjemnější analýzu zásobníku jádra a zvýraznění doby nečinnosti procesu do jeho probuzení
- ♦ možnost rozdělování událostí, které se týkají dvou procesů
- ♦ zobrazení CPU topologie u systémů využívajících NUMA

Řešení

Rozšíření jsou implementována buď jako **pluginy**, nebo jako tzv. **modifikace** zdrojového kódu aplikace, pokud nebylo možné dosáhnout požadovaných vlastností pomocí pluginů.

Některá rozšíření nechala vzniknout menším, dodatečným, nicméně samostatným vylepšením, např. přístup k vnitřním tabulkám barev, či aktualizace kódu na novou verzi Qt.



Stacklook

Plugin zobrazuje nad vybranými záznamy událostí tlačítka, která po dvojitém kliknutí zobrazí okénko se záznamem zásobníku kernelu získaného těsně po zaznamenání události. Při přejetí kurzoru myši přes tlačítko se zobrazí část ze záznamu zásobníku v horním informačním řádku aplikace.

Naps

Plugin pomocí obdélníků v grafu vizualizuje dobu nečinnosti procesu. Zároveň informuje uživatele o předchozím stavu procesu.

Record Kstack

Modifikace pro uživatelsky příjemnější zachytávání zásobníku kernelu při trasování přes Trace-cmd.

Couplebreak

Modifikace, která přidá umělé záznamy událostí do streamu dat. Umělé záznamy vznikají z existujících událostí, které spojovaly dva procesy, např. událost *sched_switch*.

Toto rozšíření tak vytváří kompatibilitu mezi pluginy, které by jinak musely měnit vlastní proces u záznamu události.

NUMA TV

Modifikace spolupracuje s programem Hwloc pro získání topologie CPU a NUMA uzlů na systému. Tato data pak zobrazuje vedle grafu událostí, kde se napojuje na CPU grafy. Uživatel pak ihned vidí NUMA uzel (a jádro), kterému CPU patří.

Závěr

- ♦ Nové GUI prvky usnadňují čitelnost a práci se záznamy zásobníků kernelu
- ♦ Lze vizualizovat dobu a předchozí stav neaktivních procesů
- ♦ Již netřeba používat dva různé programy na analýzu trasování u systémů s NUMA technologií
- ♦ Pluginy již nemusí měnit vlastní procesy událostí, čímž přestávají omezovat pluginy s jinými předpoklady pro data

Budoucnost projektu

- ♦ Vyvinuté modifikace budou integrovány do oficiálních verzí KernelSharku
- ♦ Všechny vyvinuté projekty budou prozkoumány, laděny a případně pozměněny podle kritiky
- ♦ Každý z projektů má prostor na rozšíření, například Stacklook by mohl podporovat více událostí, Couplebreak by šlo zobecnit na systém pro vytváření umělých záznamů.