io_uring: BPF controlled I/O

Wednesday, 22 September 2021 07:00 (45 minutes)

io_uring is an asynchronous I/O API crafted for efficiency, where one of the reasons for using shared rings is to reduce context switching. It got lots of features since introduction, and pushing it further we want to give away some of the control over submitting and controlling I/O to BPF, minimising the number of context switches even more.

- It should keep the number of system calls to a minimum.
- Help to lower overhead on scheduling user processes to CPUs when they have not much to do and will go to sleep briefly.
- Be an alternative to submission queue polling for latency reduction not taking as much CPU time at the same moment.

We'll go over the current design [1] and decisions, issues and plans, and hopefully it will engage a discussion and give impetus to curious minds to try it out and share ideas on how to tailor the API to fit their use cases.

[1] https://lore.kernel.org/io-uring/a83f147b-ea9d-e693-a2e9-c6ce16659749@gmail.com/T/#m31d0a2ac6e2213f912a20f5e8d88bd74f81406

I agree to abide by the anti-harassment policy

I agree

Primary author: BEGUNKOV, Pavel
Presenter: BEGUNKOV, Pavel
Session Classification: LPC Refereed Track
Track Classification: LPC Refereed Track (Closed)