Bad Storage vs. Filesystems

Tuesday, 21 September 2021 09:45 (30 minutes)

The focus of this session is on mitigating the effects of unreliable storage devices. This author works at a cloud vendor (as is fashionable now), and one of the large story arcs of the past few years has been that storage devices do not seem as reliable as we thought even a few years ago.

Specifically, I've observed that as the world moves from direct-attached spinning rust to software-defined storage on cheap devices, we increasingly must deal with large devices that corrupt data, temporarily stop responding (due to problems on the network/control plane/hypervisor/whatever), or have some odd means to request re-reads

XFS sort of mitigates some of these problems by enabling sysadmins to configure its response to certain kinds of hardware errors (mostly EIO and ENOSPC). Other filesystems lack these control knobs; how might we standardize them? The block layer has some retry capabilities, but no filesystems touch them. We don't have a general corrupted-read retry mechanism, and have not succeeded in adding one.

So what I want to know is: Who cares? Are sysadmins and users happy with the current patchwork? Do they accept the defaults? Would they like more control or better communication between layers?

I agree to abide by the anti-harassment policy

I agree

Primary author: WONG, Darrick (Oracle)

Presenter: WONG, Darrick (Oracle)

Session Classification: File Systems MC

Track Classification: File Systems MC