Silent data corruption

Dragan Simić

BarCamp Banjaluka 0x02
14. novembar 2015.

Sadržaj predavanja

- Šta je silent data corruption?
- Koji su njegovi uzroci?
- Koliko često se dešava?
- Metode prevencije
- Kako se tu uklapa RAID?
- Next-gen fajl sistemi
- Diskusija, pitanja itd.

Šta je silent data corruption?

Bilo kakva spontana izmjena u podacima smještenim unutar računara, koja nije korigovana niti primjećena od strane prisutnih mehanizama za osiguranje integriteta podataka.

Jedan od poznatijih primjera je FDIV bug u Pentium procesorima, otkriven 1994., koji je uzrokovao pogrešne rezultate za određene matematičke operacije dijeljenja u pokretnom zarezu.

Silent data corruption nije ograničen samo na one hardverske komponente kojima je primarna funkcija smještanje podataka.

Uzroci pojave silent data corruption

- Greške koje su načinjene u proizvodnom procesu integrisanih kola, kao i nepravilnosti u samoj strukturi materijala korištenih za njihovu izradu.
- Interakcija između pojedinih dijelova integrisanih kola
- Elektrostatičko pražnjenje kroz hardverske komponente
- Kosmička radijacija (visokoenergetski protoni i atomska jezgra)
- Nestabilni izvori napajanja ili loši sklopovi za regulaciju napona
- Vibracije ili jaki izvori zvuka
- Greške u firmveru hardverskih komponenti ili u softveru

Koliko često se to dešava?

- IBM: kao posljedica kosmičkog zračenja, na svakih 1 GB RAM-a dešava se sedmično jedna greška u radu.
- NetApp: tokom 41-mjesečnog posmatranja 1.5 miliona hard diskova, zabilježeno je 400,000 silent data corruptions od kojih 31,000 nije detektovao korišteni RAID podsistem.
- CERN: tokom šest mjeseci, 128 MB od posmatranih 97 PB podataka je pretrpilo trajno oštećenje.
- 2008. godine je izgubljena kontrola nad jednim Airbusom A330, uz povrede putnika, a kao posljedica uticaja kosmičkih zraka.

Metode prevencije (I)

Poboljšavanje proizvodnih procesa radi smanjenja prisustva silent data corruption, kroz kvalitetnije materijale, naprednije dizajne, detaljnije procese testiranja, itd.

Drugi pristup se orijentiše na povećanje otpornosti sistema na silent data corruption, korištenjem ECC zapisa ili kontrolnih suma, dodatnih provjera stanja internih protokola, višestrukog obavljanja istih operacija uz korištenje mehanizama glasanja, itd.

Za rekonstrukciju podataka je neophodna neka vrsta redundancije.

Metode prevencije (II)

Sa stanovišta pojave silent data curruption, slaba tačka su mjesta na kojima podaci prelaze granice između različitih podsistema, pošto se tada iznova kreiraju ECC zapisi i kontrolne sume.

Rješenje pruža tzv. *end-to-end data protection*, odn. kontrolne sume koje se generišu unutar krajnjeg “potrošača” podataka i ostaju uz podatke na svim nivoima storage steka.

End-to-end data protection može da koristi specifičan hardver (Protection Information, PI) ili da bude dio next-gen fajl sistema.

Kako se tu uklapa RAID? (I)

Hard diskovi sadrže ECC zapise na nivou sektora i obezbjeđuju na taj način provjeru integriteta podataka i njihovu rekonstrukciju, uz mogućnost remapiranja problematičnih sektora.

RAID podsistemi koriste statuse operacija dobijenih od hard diskova da odluče o integritetu podataka: ako je status “ok”, podrazumijeva se da je očuvan integritet dobijenih podataka.

Unrecoverable read errors (URE) rezultuju “error” statusima, ali dešava se i da oštećeni podaci budu vraćeni uz “ok” status.

Kako se tu uklapa RAID? (II)

RAID nivoi koji pružaju redundanciju podataka omogućavaju remapiranje i rekonstrukciju sektora koje hard diskovi nisu uspjeli da pročitaju, i pored ECC zapisa na nivou sektora.

Preventivno korištenje tzv. *data scrubbinga* omogućava ranu provjeru i rekonstrukciju podataka prije nego što obim oštećenja preraste obim redundancije koji pruža određeni RAID nivo.

Međutim, u slučaju pojave silent data corruption ovi mehanizmi ne funkcionišu bez upotrebe end-to-end data protection.

Next-gen fajl sistemi (I)

Next-gen fajl sistemi, kao što su Btrfs i ZFS, pružaju integrisanu end-to-end data protection funkcionalnost kroz generisanje i trajno čuvanje kontrolnih suma na nivou bloka.

Dodatno, next-gen fajl sistemi podržavaju i integrisanu RAID funkcionalnost, koja može da obezbijedi redundanciju podataka.

Provjera kontrolnih suma prilikom svakog čitanja podataka omogućava detekciju silent data corruption, a u kombinaciji sa obezbijeđenom redundancijom omogućava i rekonstrukciju.

Next-gen fajl sistemi (II)

Potreba za integrisanjem RAID funkcionalnosti u next-gen fajl sisteme je nastala upravo iz namjere obezbjeđivanja mogućnosti transparentne rekonstrukcije oštećenih podataka.

Pored RAID funkcionalnosti, next-gen fajl sistemi sa istim ciljem takođe integrišu i *logical volume management* funkcionalnost.

Tako je narušena tradicionalna slojevita struktura Unix storage steka, u kojoj je svaki sloj zadužen za određenu funkcionalnost, ali je upravo taj nivo apstrakcije predstavljao prepreku.

Pitanja?

Za više detalja, spisak korištenih referenci itd.
https://en.wikipedia.org/wiki/Data_corruption
https://en.wikipedia.org/wiki/Talk:Data_corruption

Autor ovog predavanja
je ujedno dao doprinos tom
i drugim srodnim člancima. :)