nedeljainformatikev2.0

Mikrokernelski operativni sistemi

David Davidović

Matematička gimnazija
Nedelja informatike v2.0

18. decembar 2015.

Operativni sistemi

- Znamo da je operativni sistem (OS) softver koji se bavi najnižim operacijama na računaru i koordinira njegovim radom
- Postojalo je mnogo operativnih sistema kroz istoriju, danas se koriste…
  - Microsoft Windows
  - Linux distribucije
  - Apple OS X i Apple iOS
  - Google Android
  - …i mnogi drugi
- Kako oni zapravo rade?

Operativni sistemi

- Dizajn i arhitektura operativnih sistema je otvoren problem u kompjuterskim naukama
- Od ranih vremena i OS-ova kao što su OS/360 i CP-67 pa do modernih istraživačkih, svaki je donosio drugačije boljitke i balans funkcionalnosti
- Operativni sistemi su takođe prisutni na *embedded* uređajima i na ponekad neočekivanim mestima
  - Aparati za kafu rade pod nekim operativni sistemom
  - *Baseband* procesor na vašem telefonu koristi neki *real-time* OS za koji verovatno niko od nas nije čuo

Kernel - srž cele stvari

- Kernel je glavna komponenta operativnog sistema
- Bavi se zadacima koji su najbliže hardveru i “najprljaviji” — direktno upravlja periferijama, omogućava *multi-tasking*, bavi se upravljanjem memorijom…
- Operativne sisteme je teško podeliti na pravi način jer mogu isuviše da variraju
- S druge strane, kerneli su nešto uže specijalizovani i možemo da razlikujemo neke grube opšteprihvaćene kategorije

- Kernele možemo (a i ne moramo) podeliti na:
  - Monolitni kerneli,
  - Mikrokerneli — o kojima ćemo danas pričati i
  - Hibridni kerneli.
- Mnogi kerneli koji postoje i koji su postojali, čak i dosta njih koji su trenutno u upotrebi, ne mogu se svrstati tačno u jednu od kategorija
- Surovost modernih procesorskih arhitektura i želja za optimizacijom, bezbednošću i sl. "naterala" je programere i softverske arhitekte da se šetaju po tankim linijama ove kategorizacije
- Ipak, ovakva podela će nam pomoći da malo bolje razumemo koncepte koji slede

Nivoi izvršavanja na procesorima

- Kako bi se postigla separacija koda koji se izvršava na procesoru, procesorske arhitekture odavno podržavaju *nivoe izvršavanja*
- Uglavnom postoje bar dva — *privilegovani* i *neprivilegovani* nivo
- Kod operativnog sistema se izvršava u privilegovanom nivou, a kod koji potiče od korisnika sistema se izvršava u neprivilegovanom
- Ovim se postiže da korisnički kod ne može pristupiti poverljivim kernelskim podacima, direktno petljati oko hardvera i slično
- Prelazak između nivoa je skup i to je glavna komplikacija

Monolitni kerneli

- Monolitne kernele karakteriše činjenica da se sav kod njihovih komponenti — drajvera, mrežnih biblioteka i raznih drugih delova izvršava zajedno u privilegovanom nivou
- Ovo je dobro jer se znatno smanjuje broj *kontekstnih prelaza* odnosno prelaza između nivoa izvršavanja
- Ovo nije dobro jer je sigurnost i stabilnost sistema kompromitovana: bag ili propust u drajveru utiče na ceo sistem!
- Komponente često komuniciraju deljenom memorijom
- Primeri: Linux, DOS…

Monolitni kerneli — ilustracija

Hibridni kerneli

- Hibridni kerneli su *catch-all* grupa — mešavina monolitnih i mikrokernela, sa raznim balansom između te dve arhitekture
- Često su kerneli unutar popularnih i sveprisutnih OS-ova hibridni jer im to omogućuje da iz praktičnih razloga uzmu "bolje iz obe varijante"
- Primeri: Windows NT (kernel unutar Microsoft Windows-a), XNU (kernel unutar Mac OS X-a i iOS-a), …

Mikrokerneli

- Kod mikrokernela se u privilegovanom procesorskom nivou nalaze samo najosnovnije funkcije
- Drajveri, mrežni protokoli, fajl sistemi, memorijski upravljači i sl. su najobičniji procesi kojima kernel dozvoljava neke interesantne stvari
- Sistem je stabilniji — krešovanje audio drajvera ne znači da se računar restartuje i prekine me u gledanju novog spota Sandre Afrike
- Sistem je sigurniji — bezbednosni propusti su većinom mnogo manje kritičnosti
- Sistem je sporiji — mnogo kontekstnih prelaza, nema deljenja memorije pa se moraju slati poruke između komponenti, itd.

Mikrokerneli - ilustracija

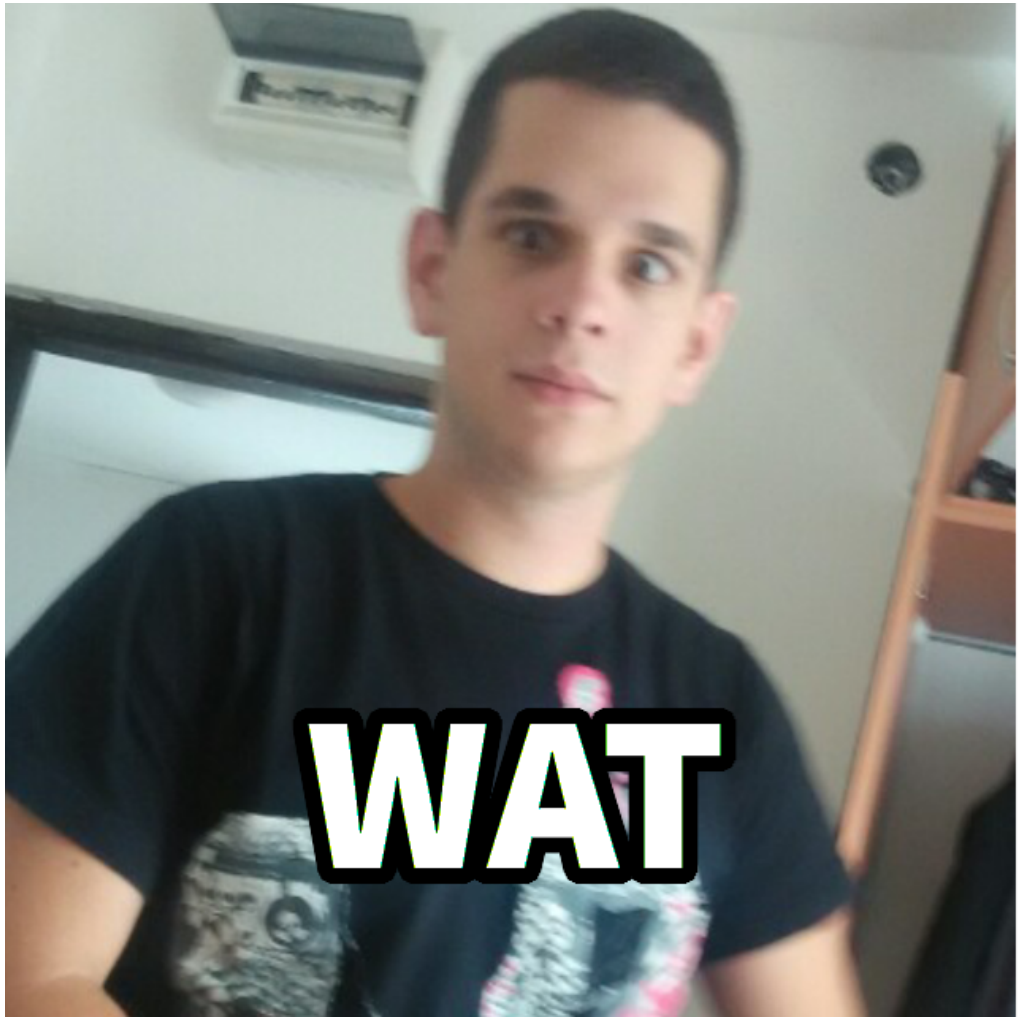
Sigurnost

- Generalno, mikrokerneli su sigurniji od svojih monolitnih sunarodnika
- Zbog vrlo snažne izolacije između raznih komponenti, propusti u nekima od njih imaju ograničen doseg
- Koordinator celog sistema, tzv. *mikrokernelski server* je jedina komponenta koja mora biti bez propusta kako bi se osigurala vrlo bezbedna arhitektura
- Ovo omogućuje nekim zaluđenicima da formalno dokažu ispravnost mikrokernela ili njegovog interfejsa

CVE-2010-3904

- Kao primer uzimamo jedan od mnogobrojnih bagova u kodu Linux kernela koji omogućavaju napadaču da uradi neke ružne stvari
- Implementacija mrežnog protokola RDS u Linux kernelu nije proveravala datu destinacionu adresu prilikom kopiranja neke memorije
- Bilo je trivijalno dati RDS podsistemu adresu kernelske memorije i naterati ga da upiše proizvoljan kod u kernelski memorijski prostor
- Time napadač koji ima bilo kakav pristup Linux sistemu može da preuzme potpunu kontrolu nad njim

CVE-2010-3904



CVE-2010-3904

- Pod mikrokernelskom arhitekturom, ovakav propust u opskurnom mrežnom protokolu ne bi mogao da dovede do katastrofalne kompromitacije sistema
- Nažalost, u pitanju nije usamljen primer:
  - CVE-2013-2094 je koristio bag u profajleru ugrađenom u Linux kernel za dobijanje *root* pristupa,
  - CVE-2015-1328 je u iste svrhe koristio bag u kodu za OverlayFS fajl sistem,
  - CVE-2015-1318 je dozvoljavao korisnicima da kompromituju sistem pomoću *crash reporting* alata zvanog *Apport*,
  - Na slične bisere u Windows NT-u ne vredi trošiti reči.

Stabilnost

- Iz sličnih razloga, stabilnost sistema je povećana
- Mikrokernelski server restartuje sistemske procese koji krešuju, što znači da, na primer, u slučaju baga u nekom drajveru ne dolazi do *kernel panic*, *blue screen of death* situacije ili smrzavanja smartfonova
- Ovo je posebno bitno za servere i slične sisteme koji moraju biti stalno dostupni
- Andrew S. Tanenbaum, poznati profesor kompjuterskih nauka i autor MINIX operativnog sistema, kritikovao je nestabilnost ovakvih monolitnih arhitektura rečenicom “kada je poslednji put Vaš TV uređaj krešovao i morao da se restartuje?”

Elegancija

- Iako diskutabilne važnosti, valja napomenuti da su mikrokernelski dizajnovi uglavnom mnogo elegantniji od monolitnih i hibridnih
- Jasna separacija komponenti omogućuje mnoge estetski prijatne nivoe apstrakcije i enkapsulacije koje nije moguće postići u alternativnim arhitekturama
- Dobar primer je upravljanje memorijom u L3/L4 porodici mikrokernela koju je dizajnirao Jochen Liedtke
  - Kroz koncept *rekurzivnih memorijskih strana* i formalno definisanim operacijama nad njima, Liedtke je dizajnirao sistem kojim se razni specijalizovani upravljači memorijom mogu pokrenuti kao sistemski procesi bez gubitka sigurnosti

Brzina

- Nažalost, OS-ovi bazirani na mikrokernelima su kroz istoriju bili mnogo sporiji od hibridnih i monolitnih
- Činjenica da se skoro sve kernelske funkcije sem najbitnijih izvršavaju u korisničkom prostoru povlači i veliki broj kontekstnih prelaza (koji traju dosta na procesorima)
- Mikrokernele često dizajniraju zaluđenici za lepom, čistom arhitekturom i performanse im nisu prioritet
- Pritom, zbog nesrećnog toka istorije, procesorske arhitekture su optimizovane u pravcu boljeg rada sa monolitnim i hibridnim kernelima

Nepraktičnost

- Iako imaju svojih prednosti, mikrokerneli su mahom nepraktični za većinu *general-purpose* upotreba
- Dizajn *message passing* šema za komunikaciju među komponentama je nezahvalan posao koji usporava razvoj
- Većina prednosti koje se dobijaju nisu dovoljne da opravdaju razvoj novog, brzog, modernog i praktično upotrebljivog mikrokernela
- Monolitni kerneli su toliko sveprisutni da se često u nov sistemski softver "ugrađuju" određene pretpostavke koje otežavaju moguću adopciju mikrokernela

Nerasprostranjenost

- Najzad, mikrokerneli pate i od *kokoška-jaje situacije*
  - Šta me gledate, pa nisam ja smislio taj naziv
- Nisu mnogo rasprostranjeni i zbog toga nema mnogo inicijative da se radi na unapređivanju, a unapređivanje je ono što je potrebno kako bi mogli da pariraju trenutnim monolitnim i hibridnim rešenjima
- Većina mikrokernela koji postoje su ili samo istorijski važni, ili su napravljeni kao istraživački projekti kojima nije ni bio cilj da postanu spremni za generalno korišćenje
- Iz ovog razloga, cena koju je potrebno platiti za dodatnu stabilnost i sigurnost je prevelika i nema ekonomskog ni zdravorazumnog smisla

L3/L4 porodica

- L3 i L4 su porodice mikrokernela od kojih postoje nekoliko varijanti
- Imaju bogatu istoriju i često se kerneli zasnovani na njima koriste unutar nekih specijalizovanih industrija i upotreba
- seL4 iz L4 porodice je jedinstven po tome što je u pitanju jedan od retkih praktičnih mikrokernela koji su pritom i formalno dokazani
  - "Formalno dokazani" znači da je matematički dokazano da su interfejsi kernela korektni i da garantuju bezbednost i stabilnost
  - Iz ovog razloga je seL4 našao primenu u industrijama poput vojnih primena i razvoja bespilotnih letelica
- Ove porodice su takođe uvele i inovativne koncepte, posebno one koji su poboljšale brzinu današnjih mikrokernela

Mach

- Mach je stari mikrokernel razvijen od strane Carnegie Mellon univerziteta davnih osamdesetih godina prošlog veka
- Značajan je zbog toga što je njegova arhitektura oformila arhitekturu nekih današnjih *general-purpose* hibridnih i mikrokernela
  - XNU, hibridni kernel koji pokreće sve Apple uređaje je baziran na Mach kernelu, s tim što se u njemu korisnički procesi takođe izvršavaju u kernelskom prostoru što efektivno negira većinu prednosti koje mikrokernelska arhitektura pruža ali nudi značajne dobitke u brzini
  - GNU Hurd, Unix-kompatibilni operativni sistem koji je skoro dobio svoje prve verzije koje mogu praktično da se koriste kao serverski i desktop sistemi, baziran je na Mach-u

Tekuća istraživanja u ovoj oblasti

- Mnogi istraživački mikrokerneli pokazali su da imamo načine da skoro neutrališemo performansne probleme koje ova arhitektura predstavlja
- Xen, softver koji se koristi na velikom broju servera, koristi mikrokernelski dizajn, i pokazano je da u nekim situacijama adaptirani (“paravirtuelizovani”) Linux kernel pod Xen-om može da neke operacije izvršava brže nego Linux koji radi direktno na hardveru!
- Ipak, verovatnoća da će se mikrokernelski sistemi koristiti u skorijoj budućnosti je nepostojeća
  - Cela industrija, posebno u vezi sa starim i inertnim konceptima kao što su operativni sistemi, pati od velikog i teškog nasleđa u vidu starih sistema koji bi mogli bolje da se “od nule” napišu

NK

- NK (*novi kernel*) je ideja koju sam izložio u svom maturskom radu — mikrokernelska arhitektura koja se oslanja na softver, ne hardver da pruži izolaciju između komponenti sistema
- Izvršavanjem koda koji je u mašinskom obliku, koji može da pokuša da uradi bilo šta na sistemu, nije moguće izbeći da sam procesor, kroz nivoe izvršavanja, mora to da kontroliše
- Međutim, čuvanjem koda u nekom međuobliku u kome se takve operacije ne mogu predstaviti, možemo da ga izvršimo u kernelskom nivou tako da i dalje imamo sve garancije sigurnosti
- Glavna ideja je brutalno smanjiti performansne probleme sa kontekstnim prelazima kroz softverski, a ne hardverski održanu bezbednost

Nasleđe i paraliza

- *Nasleđe i paraliza* su reči kojima je u jednom humorističnom govoru Gary Bernhardt (neki lik koji drži humoristične govore) sumarizovao stanje naših računarskih sistema
- Nasleđe — jer imamo stare sisteme i dalje u upotrebi, još starije sisteme sa kojima moramo da održavamo kompatibilnost, i sl.
- Paraliza — jer zbog ovih zahteva za kompatibilnošću i interoperacijom sa njima ne možemo da uvedemo velike promene
- U pitanju je veliki problem, izazvan brzinom kojom je industrija eksplodirala i zbog koje nije imala vremena za promenu mnogih starih arhitekturalnih odluka koje sada čine mnogo više štete nego koristi

Mikrokerneli su zapravo kul

- Većina problema zbog kojih danas mikrokernelski operativni sistemi nisu upotrebi je odavno nestala
  - Procesori su sada brži
  - Ljudi su sada manje glupi
- Kada bi ovakvi praktično upotrebljivi OS-ovi postojali, mnogo problema koje trenutno imamo bi nestala ili bila smanjena
  - Manje kritičnih sigurnosnih propusta koji mogu da izbuše svaki računar koji koristi određeni OS
  - Manje neobjašnjivih odbijanja saradnje od strane računara baš kada treba da za 3 sata završite taj sastav (ili prezentaciju za Nedelju informatike)
  - Sistemi koji se lakše održavaju i olakšavaju život programerima na duge staze

Mikrokerneli su zapravo kul

- Nažalost, kao što je već spomenuto, investicija koju imamo u trenutnim sistemima nam ne dozvoljava da ozbiljno razmišljamo o njihovoj promeni zbog haosa koji bi usledio
- NK je mikrokernel za početak namenjen smartfonovima, gde je taj problem manje izražen i gde je lakše razviti takvo rešenje i neprimetno ga poturiti korisnicima
- Načelno, to je svakako nešto što ima jako malu šansu da uspe iz razloga koje smo spomenuli
- Ipak, ako dođete do bilo kakve inovacije koja funkcioniše bolje i elegantnija je od postojećih rešenja, najveća usluga koju možete da učinite sebi i drugima je da je predstavite i progurate bez obzira na to koliko to suludo zvuči
  - Postojeća rešenja su uglavnom tako i počela :)

Pitanja?

- Hvala na pažnji!