Maskininlärning för effektivitets hantering inom flerkärniga processorer

Innehållsförteckning

1. Inledning

Transistorantalet i processorer har ökat mycket det senaste årtiondet och därmed har densiteten ökat för processorer, men nyligen har ökningen avtagit på grund av att transistorstorleken på nya processorer redan är så liten att det har blivit svårt att minska transistorerna mera. Av det här uppkommer det problem för att öka på prestandan varje år och för att överkomma detta problem har processortillverkarna ökat energiförbrukningen till processerna som ökar på klockfrekvensen och därmed ökar på prestandan. På grund av ökningen på energiförbrukningen har det uppkommit ett nytt problem för nedkylningen för processorer. Dagens stationära datorer kan till och med använda över 1000W och själva processorn kan använda ca. 300W beroende på modellen. Konsumenter för stationära datorer har ens viss frihet för olika nedkylnings alternativ och det har uppkommit att det är ändå relativt svårt att hålla temperaturen för processorer under kontroll. Bärbar dator har det ännu svårare fastän de använder märkbart mindre energi för att på en bärbar dator har man inte väldigt mycket plats för ordentlig ned förkylnings lösning. Fastän processorer för bärbara datorer använder mindre energi än processorer för stationära datorer är det ännu viktigare att ha en bra energihanterings protokoll för att en bärbar dator använder ett batteri för energin. För att hålla en processor under rätta temperatur, energi och prestanda skala så måste processorn justera på spänningen och effekten, det här kontrolleras av den traditionella energihanterings metoden. När en processor ska ha en högre prestanda måste dess klockfrekvens öka och för det behöver processorn mera spänning och effekt som ökar på temperaturen. Ökningen på temperaturen minskar på effektiviteten, det menar att processorn måste använda mera spänning som ökar igen på temperaturen. Processorer har inbyggda säkerhets mätningar som stoppar processorer att nå en viss temperatur, spänning och effekt. Inom dessa säkerhets ramar kan processorer justera sig så att den har den högsta klockfrekvens beroende på belastningen. Traditionella energihanterings metoden som tidigare nämnts har kontrollen över detta och bestämmer när och hur mycket skall klockfrekvensen öka på för en viss belastning och för det ska spänningen

och effekten ökas också. Den Traditionella energihanterings metoden har visats börja bli föråldrad för nästa generations processorer på grund av den höga densiteten och för att den traditionella energihanterings metodens implementation har vissa flaskhalsar som man inte kan överkomma.

1.1 Traditionell energi och temperatur hantering

Den traditionell energihanterings och temperatur kan skalas ner till den Dynamiska spännings- och frekvensskalning metoden (eng. Dynamic voltage and frequency scaling) (DVFS). Den här metoden har använts redan i många år och har fungerat såväl som bra men för nya moderna processorer med flera kärnor och högre densitet har metoden börjat visa sig vara föråldrad. Problemet med DVFS är att metoden inte är implementerad till en specifik processor utan tillverkarna implementerar en statistisk modell för en viss processormodell och därmed kan inte DVFS fungera på sin maximala förmåga. Processortillverkarna kan inte veta vilka kylningsmetoder kommer användas för de processorer som de tillverkar, kylningsmetoden har en stor andel hur mycket en processor kan öka på sin klockfrekvens och därmed en bättre kylningsmetod hjälper med att förbättra energieffektiviteten, desto lägre temperatur desto mindre spänning behöver en processor för samma klockfrekvens. Som exempel, kan två samma modells processor köras med olika kylnings lösningar med helt olika resultat till exempel processor 1 kan ha en temperatur på 80C grader och processor 2 kan ha en temperatur på 60C garder. Det här är orsaken för att DVFS måste implementeras med en statistisk modell som iakttar den värsta scenariot och det här den största flaskhalsen för DVFS. En annan flaskhals för DVFS är att under de senaste årtionden har spänningen minskat för kärnorna men effekten har ökats. Av det minskade kärnspänningen förekommer läckageeffekt. Med mindre kärnspänning har DVFS ett mindre fönster för att justera på spänningsskalan. Kiseltransistorer har en minimumbegränsning av kärnspänningen som är 0,7V och maximum runt 1.1V till 1.4V, det här försvårar processen på att spara energi med DVFS metoden för att skalan mellan minimum och maximum spänningen är relativt liten. Äldre kiseltransistorer har en mycket högre skala på minimum och maximum spänningen och därför har DVFS

metoden kunnat spara mera energi för de äldre processorerna. Hur DVFS sparar energi med att sänka på klockfrekvensen kan visas med P = C * f * V^2 + P statisk var C är kapacitansen för transistorgrindarna, f som är driftfrekvensen och V som är matningsspänningen.[1]
https://www.usenix.org/legacy/events/hotpower/tech/full_papers/LeSueur.pdf

För att DVFS metoden lider av de tidigare nämnda flaskhalsarna så kommer DVFS börja bli föråldrad för de kommande processorgenerationerna. De här flaskhalsarna förekommer på x86 arkitekturens processorer som tillverkats av AMD och Intel, ARM processorer kan möjligen också använda DVFS för effektivts hantering men lider inte troligen lika mycket av flaskhalsarna. Ett sätt över komma statistiska implementations problemet på DVFS är med maskininlärning (ML).

1.2 Historia av varför effekttätheten har ökat

För att förstå varför en bra energi och temperaturs hanteringsmetod är viktigt måste vi förstå lite om historian bakom vad har hänt med utvecklingen inom processorer. När processortillverkarna utvecklar en ny generation så ökade man oftast på transistor antalet och ibland på kärnorna. Med mera transistorer ökar effekten som ökar på temperaturen. Processortillverkarna märkte att när de utveckla en ny generation med ny processnod kunde de minska på spänningen som gjorde det möjligt att temperaturen ökade relativt lite för att minskningen på spänningen drog ner på temperaturen fast än effekten hade ökat. Det här möjligt gjorde att konsumenter såg en relativ lite höjning per generation. Senaste året har det här inte mera fungerat lika bra för att processortillverkarna har i princip maxat ut denna effekt till spännings förhållandet och kan inte mera minska på spänningen tillräckligt mycket för att kompensera effekttätheten. Utan ett förskott i processnod så kommer nya generationer i framtiden också ha höga temperaturer. Ett annat problem som har uppkommit med ökningen på effekttätheten är att värmespridaren på processorer fungerar nästan som en värmeisolator som försämrar värmeledningsförmågan. Det här förekommer inte alltid men med rätta kombinationer på processor storlek på (DIE) och på kylningsmetoden

så kan värmespridaren värka som en värmeisolator. Värmespridaren är en viktig del för processorn för att de skyddar själva (DIEn) när man byter kylpasta på sin processor och därför kommer inte processortillverkare ta bort det fast än den kan agera som en värmeisolator i vissa fall.

2. Grundläggande information

En processor har olika tekniker och kontroller för att hantera och kontrollera olika värden på processorn, så att processorn är stabil, presterar och effektiv. I detta kapitel presenteras grundläggande information av dessa kontroller och tekniker.

2.1 Effekt- och energiförbrukning

När en processor gör arbete ökar effektförbrukningen (Watt W) som producerar värme, effektförbrukningen förändras över tiden beroende på flera olika variabler. Olika beräknings uppgifter för en processor påverkar på effektförbrukningen beroende på hur krävande uppgiften är. Arkitekturen och processnoden påverkar också på effektförbrukningen, det här kan man inte påverka på utan fastlagda fysiskt. Frekvensen och spänningen på kärnorna ökar också på effektförbrukningen och strömläge så som aktiv, inaktiv och viloläge. Temperaturen påverkar också på effektförbrukningen, desto högre temperatur desto mindre effektivitet. Effektförbrukningen mäts momentant med enheten Watt (W), För att mäta energiförbrukningen måste vi mäta effekten över en tids period med enheten Joule (J) eller Wattsekund (Ws).

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8509120&tag=1

2.2 Temperatur hantering (DTM och TDP)

Temperaturen på en processor ökar när effektförbrukningen ökar fastän effektförbrukningen kan ändras ögonblickligen så ökar inte temperaturen på samma sätt. För att hålla en processor i rätta temperaturer måste man installera en kylningslösning som har en högre termisk kapacitans än själva processorn. Med en kylningslösning ökar temperaturen så väl som jämnt beroende på kylningslösningens förmåga att avleda värme. Om kylningslösning är värmemättad och kan inte avleda värme tillräckligt snabbt kan det förorsaka permanenta skador för processorn om temperaturen ökar över kritiska temperatur värdet som oftast är runt 100 Celsius grader. Processortillverkarna ger oftast en värde på termisk designeffekt (thermal design power

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8509120&tag=1

(TDP)) som ges med enheten Watt. Det här hjälper med att välja en rätt kylningslösning för processorn. TDP är definierad som ”den högsta förväntade hållbara effekten under körning effektintensiva verkliga applikationer”. Beräkningen för TDP varier på tillverkare och oftast är TDP värdet inte alls verkliga för konsumenter som kan kategoriseras som kraftanvändare. Fast än processor har en rätt kylningslösning enligt TDP kan kylningslösningen bli värme mättad för att TDP beaktar inte den maximala effekten som kan förbrukas. I denna fal om en kylningslösning blir värme mättad och kan inte avleda värme tillräckligt snabbt så har tillverkare implementerat en dynamisk värmehantering (dynamic thermal management (DTM)). DTM viktigaste uppgift är att stänga av kärnor, reducera spänningen och frekvensen när den kritiska temperaturen har nåts så att processorn inte skadar sig.

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8509120&tag=1

2.3 Dynamisk energihantering (DPM)

Dynamisk energihantering (DPM) uppgifter är att dynamiskt byta effekttillstånd på enskilda kärnor. Kärnorna kan ha olika tillstånd så som exekverings/aktiv läge eller till lågströmsläge till exempel inaktiv/klockstyrd, viloläge och strömstyrd etc. De olika lågströmslägen är beroende på tillverkare och processormodell, de olika lägen har en

direkt påverkan på energiförbrukningen men med att byta tillstånd förekommer det latens som påverkar negativt på prestanda. Oönskade situationer som kan förekomma när DPM byter tillstånd på kärnorna är till exempel när en kärna är satt till viloläge och genast efter det bli väckt. Det här förorsakar latens och extra energiförbrukning. Ett annat exempel kan vara när en kärna är satt på inaktiv (tomgång) tillstånd och i denna period förekommer det inga serviceförfrågningar. På denna period har nu kärnan slösat energi med att inte byt tillstånd till ett lågströmsläge. Det finns tre huvudsakliga DPM metoder Fixed time-out, Adaptive time out och predicitve techniques. Fixed time-out metoden stänger av kärnor efter en de har varit inaktiva (tomgång) en viss tid som är fixerat. Adaptiv time-out metoden är mer effektiv än Fixed time-out, Adaptiv time-out anpassar time-out tid enligt historik. Predictive techniques väntar inte på att time-out tiden har upphört utan stänger kärnorna genast när de har byt tillstånd till inaktiv (tomgång) om inaktiv tiden har förutspåtts att bli tillräckligt lång så att det inte förekommer negativa påverkningar.

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8509120&tag=1

2.4 Dynamisk spänning och frekvensskalning (DVFS)

Dynamisk spänning och frekvensskalning kontrollerar spänningen och frekvensen på kärnorna i en processor. Spänningen på en kärna och frekvensen är relaterad till varandra, för högre frekvens måst vi höja på spänningen på kärnan. Relationen mellan spänningen och frekvensen kan visas som fstable = k * ((Vdd-Vth)^2/Vdd) var fstable är den maximala stabila frekvensen för en kärna, Vdd är matningsspänningen, Vth är tröskeln för spänningen för en viss teknologi (node?) och k är arkitekturberoende passningsfaktor. För en given Vdd som körs med lägre frekvens än fstable är effekt/spännings ineffektivt. Det finns olika variationer på DVFS för olika processorer och modeller. Olika variationer kan se ut som följande.

1. Global spänningen och frekvensskalning: Alla kärnor i processorn delar gemensam spänning och frekvens.

2. Global spännings skalning: Alla kärnor i processorn delar en gemensam spänning men enskilda kärnor kan ändra på sina frekvens oberoende.

3. Spänning och frekvens öar/kluster: Olika grupper av kärnor delar en gemensam spänning och frekvens.

4. Spännings öar/kluster: Olika grupper av kärnor delar en gemensam spänning, men enskilda kärnor kan ändra sina frekvenser oberoende.

5. Per kärnspänning och frekvensskalning: Spänningen och frekvensen för alla kärnor väljs individuellt.

Energiförbrukningen påverkas kvadratiskt av spänningen, därför är det väldigt viktigt att ha en optimerad spänning men en optimerad spänning kan i vissa fall vara väldigt komplexa att implementera. Per-kärnspänning och frekvensskalning är den effektivaste sättet för att kärnorna får sin individuella spänning som är den minimumspänning för en viss frekvens. Men per-kärnspänning och frekvensskalning lider av extra komplexitet. I andra sidan Global spännings skalning är den ineffektivaste sättet av de fem metoderna men har den enklaste implementationen.

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8509120&tag=1

Ett problem med DVFS är att den inte är kalibrerad till en specifik processor utan skapas på en statistisk modell för att olika processorer av samma modell SoC är inte identiska utan de har unika egenskaper. Till exempel om vi skulle jämföra två processorer av samma modell så kan processor 1 vara stabil på frekvensen 5GHZ med 1.3 Volt medan processor 2 med en sämre kvalitet på SoC kan inte alls köras på 5GHZ utan måste köras på 4.7GHZ med samma spänning. Ett annat exempel kan vara att processor 2 kan köras på 5GHZ men den måst ha 1.35V för att var stabil. Av den här orsaken har DVFS en ganska stor marginal för bästa prestanda, marginalen är ungefär 10-20%. Ett annat fenomen som påverkar på DVFS prestanda är temperaturen på processorn som kan variera väldigt mycket per användare. Det här är också en orsak för att DVFS implementeras statistiskt och för den höga marginalen.

https://www.extremetech.com/gaming/203898-amd-details-new-power-efficiency-improvements-update-on-25x20-project

Ett annat fenomen som har uppkommit per ny generation är att kärna spänningen har minskat när transistor storleken har minskat. De här menar att gamla processorer med hög kärnspänning har en större skala för att minska på spänningen t.ex. 1-5V. Moderna processorer har en kärnspänning from 0.7-1.4V beroende på processorn och modellen men oftast runt dessa värden. Skalan på minimum och maximumspänning på moderna processorer är väldigt mycket mindre än på de gamla, det här menar att DVFS har det svårare att minska på spänningen på en viss frekvens på de moderna processorerna på grund av den minskade skalan på kärnspänningen. Av dessa orsaker har prestandan på DVFS minskat och för kommande processorgenerationer kommer vi troligen se att tillverkare kommer överge DVFS för något annat system. De här nackdelarna påverkar först och främst på processorer från Intel och AMD och processorer som är tillverkade för stationära datorer och servrar.

https://www.usenix.org/legacy/events/hotpower/tech/full_papers/LeSueur.pdf

3. Maskininlärning

Det finns olika typer av maskininlärning som kan implementeras för energi-, effekt- och värmehantering. Detta kapitel går igenom olika typer av maskininlärning som möjligt vis skulle kunna bli använd för energi-, effekt och värmehantering. De olika typerna av maskininlärning som presenteras är: Övervakat inlärning (eng. Supervised Learning), Oövervakad inlärning (eng. Unsupervised Learning), Förstärknings inlärning (eng. Reinforcement Learning), Q-inlärning, (eng. Q-Learning) och TD(λ)-inlärning (eng. TD(λ)-Learning).

3.1 Övervakad inlärning

Övervakad inlärning fungerar på en vis tränings data mängd, oftast består data mängden av vektor inmatningar (eng. input). I övervakad inlärning har man också utmatnings värden (eng. output) som fungerar som övervakningssignaler. En

övervakad inlärnings algoritm analyserar tränings data mängden och övervakningssignalerna, därefter kan den producera estimerad funktion där funktionen kan användas för att kartlägga nya exemplen. Man kan beskriva algoritmen med given tränings samplar N, för exempel {(x1, y1), (x2, y2),…,(xN, yN)} var xi utgör funktionsvektorn för ith(check) inmatnings exempel. På motsvarande sätt kan man märka yi (ith) var vi får en model F där F(X) = Y eller F : X → Y baserat på inmatningen X med motsvarande Y.

4. Effekthantering med Övervakad inlärning

Enligt fynden på forskningen i [??] är effekthanteringen förbättrad med övervakad inlärning.

6. Referenser

[1] EL. Sueur, G. Heiser, ”Dynamic Voltage and Frequency Scaling: The Laws of Diminishing Returns, https://www.usenix.org/legacy/events/hotpower/tech/full_papers/LeSueur.pdf