

12.2 File organization and access

Wednesday, April 29, 2020 3:38 PM

I denne delen bruker vi ordet filorganisering i referanse til den logiske strukturering av poster på den måten de aksesseres. Den fysiske organisering av filer i sekundærminnet avhenger av blokkeringsstrategier og fil-tildelingsstrategier.

Når man velger måten å organisere filer på er det flere kriterier som er viktig:

- Kort aksesstid
- Enkel å oppdatere
- Økonomien ved lagring
- Enkel vedlikeholding
- Pålitelighet

Vi ser nærmere på fem ulike implementasjoner av filorganisering.

Haugfil / Ustrukturert fil

Data samles kronologisk i rekkefølgen de ankommer. Postene kan ha varierende lengde og varierende antall felter. Hensikten er å akkumulere en rekke data og lagre det. Å gjøre søk i ustrukturerte filer er veldig ineffektivt. Disse brukes dermed når det er vanskelig å organisere data. Fordelen er at de er plasseeffektive og lette å oppdatere – likevel brukes det ikke så mye.

Sekvensiell fil

Sekvensielle filer er den vanligste formen for filer. De har en bestemt lengde på poster med et bestemt antall felter i en fast rekkefølge og baseres på nøkkeltfelt. Sekvensielle søk er lite effektivt, og ikke optimalt hvis man skal gjøre enkeltsøk på nøkkel eller andre kriterier.

Indeksert sekvensiell fil

Den indekserte sekvensielle filen overkommer en rekke av ulempene med den sekvensielle filen ved å innføre en indeks til filen for å støtte tilfeldig aksess (*random access*) og en overflow-fil med peker til. For å finne et spesifikt felt søkes det i indeksfilen for å finne den høyeste nøkkelverdien som er lik eller mer enn den ønskede nøkkelverdien. Søket fortsetter i hovedfilen i lokasjonen til pekeren. Kan anvendes med flere nivåer av indeksering for mer effektivitet ved aksess.

Indeksert fil

Dersom en ønsker å søke basert på noe annet enn nøkkeltfelt er indekserte filer uegnet. For å oppnå fleksibiliteten kan man strukturere på flere indekser. I en indeksert fil aksesseres poster etter indeksene sine. Resultatet er at det ikke er begrensninger i plasseringen av posten, så lenge pekeren til minst én indeks peker på posten. I tillegg kan postene være av varierende lengde. Indekserte filer brukes mest i applikasjoner hvor punktligheten av informasjon er kritisk og hvor data sjeldent prosesseres uttømmende. Eksempler er reservasjonssystemene hos flyselskaper.

Direkte eller hashet fil

En direkte eller hashet fil utnytter evnen disker har til å direkte aksessere en blokk med kjent adresse. Det kreves derfor et nøkkeltfelt, men sekvensiell ordning benyttes ikke. Direkte filer er ofte brukt hvor veldig rask aksess er nødvendig, ved direkte aksess med poster av fast lengde og hvor poster alltid aksesseres en om gangen. Eksempler er navnelister og pris-tabeller.