Återskapningsprotokoll — nvz

Varje fil md5-summerad före packning och efter uppackning, på 12,129 filer och 2,33 GiB över 9 korpusar. Utfall: samtliga manifest identiska, och `md5sum -c` bekräftar.

2026-11-17 22:06 · Ryzen AI 9 HX 470, 24 trådar, NVMe · nvz 2.0 (c++) · läge `liten` (standard) · manifesten ligger i [`resultat/protokoll/`](../resultat/protokoll/) och protokollet genereras av [`prov/protokoll.py`](../prov/protokoll.py)

Metod

För varje korpus, i denna ordning:

1. **Manifest före:** md5 beräknas för varje fil och skrivs i `md5sum`-format till `<korpus>-fore.md5`.
2. **Packning** med `nvz packa` i standardläget `liten`; arkivets storlek och md5 bokförs.
3. **Uppackning** med `nvz packa-upp` till en tom katalog.
4. **Manifest efter:** samma beräkning på det uppackade trädet, till `<korpus>-efter.md5`.
5. **Jämförelse** rad för rad — och därtill körs systemets `md5sum -c` med *före*-manifestet mot det *uppackade* trädet, så att kontrollen görs av ett oberoende verktyg och inte av samma kod som skrev manifesten.

Manifestens egen md5 (en summa över alla filsummor) är det som visas i tabellen: är den lika före och efter är varje enskild fil lika, eftersom en enda avvikande fil ändrar manifestsumman.

Protokoll

korpus	filer	rå MiB	arkiv MiB	manifest-md5 före	manifest-md5 efter	utfall	`md5sum -c`
Programkod	5 974	55 3	7 3	`b561ff3d28d8e5c9e95ec30b9e6dc989`	`b561ff3d28d8e5c9e95ec30b9e6dc989`	**LIKA**	ja
Bilder (GIF)	105	27 4	14 2	`8e05472c70bc020586e3785a888b648b`	`8e05472c70bc020586e3785a888b648b`	**LIKA**	ja
Foton (JPEG)	1 143	408 5	345 1	`8989c1473571a2413c01ddbe118f5779`	`8989c1473571a2413c01ddbe118f5779`	**LIKA**	ja
Arkiv (zip/tar.gz)	133	391 8	242 5	`53081f8aa7bd04c35ca0f8b8fc5b2623`	`53081f8aa7bd04c35ca0f8b8fc5b2623`	**LIKA**	ja
Loggfiler	76	238 3	10 0	`b057f895ed43459cedb48d3a6b774e64`	`b057f895ed43459cedb48d3a6b774e64`	**LIKA**	ja
Bilder (PNG)	4 618	431 1	208 8	`4f99ac4f9efafab19af53eedf7dda478`	`4f99ac4f9efafab19af53eedf7dda478`	**LIKA**	ja
Film	67	536 7	532 0	`c4b518236e6e6fd1afb28e80a406082f`	`c4b518236e6e6fd1afb28e80a406082f`	**LIKA**	ja
Silesia	12	202 1	43 5	`04228513e32fe4be7f7fecda5972525d`	`04228513e32fe4be7f7fecda5972525d`	**LIKA**	ja
enwik8	1	95 4	20 4	`376ef224b8fb40ef717dc7286ac96058`	`376ef224b8fb40ef717dc7286ac96058`	**LIKA**	ja

Stickprov ur manifesten

Manifesten är vanliga `md5sum`-filer — här är början av tre, ordagrant:

Programkod — tre första raderna ur [`kod-fore.md5`](../resultat/protokoll/kod-fore.md5):

```
3cf8271b88dcad0498e0bafe268b3bd7  aix-pro/hemsogrisen-crm/.claude/settings.local.json
6a11b4fc03f849188fa66be3c736d3ea  aix-pro/hemsogrisen-crm/README.md
ffaea4222d43e7f9b929c3c082cf0e4e  aix-pro/hemsogrisen-crm/data/46elks-vakt.json
```

Foton (JPEG) — tre första raderna ur [`foto-fore.md5`](../resultat/protokoll/foto-fore.md5):

```
805dbf566034e9dd78b69d1193f73990  lc-klipp/tiktok-7186272001042943237.jpg
cd950c480a6553183cbb8b3b79a0e99b  lc-klipp/tiktok-7194110055736806662.jpg
f8b3ca51e93654772114c91a7fdb9e7d  lc-klipp/tiktok-7199292449720716549.jpg
```

Loggfiler — tre första raderna ur [`logg-fore.md5`](../resultat/protokoll/logg-fore.md5):

```
2ef8ea5eaef6d8d171834f8521bf9d8a  app/chrome-native-host.log
81d990cedd1e5ed127f942ac4669cbda  app/claude.ai-web.log
a021ad2c4bb5d4ac68ed27921d4e9d5f  app/cowork_vm_node.log
```

Kontrollera själv

Hela kedjan går att köra om på egen maskin:

```
git clone https://github.com/hugowikstrom/nvz && cd nvz && make
python3 korpus/bygg-korpus.py          # eller peka på egna kataloger
python3 prov/protokoll.py
```

Eller granska bara ett befintligt manifest mot ett uppackat träd:

```
nvz packa min-katalog arkiv.nv
nvz packa-upp arkiv.nv /tmp/ut
cd /tmp/ut/min-katalog && md5sum -c /sökväg/till/manifest-fore.md5
```

Noteringar

- md5 används här för att det är formatet folk känner igen och kan köra

med `md5sum -c`. Packarens egen uppackningskontroll jämför därutöver blake2b-128 mot arkivregistret för varje fil, och den tidigare rapporten verifierar även rättigheter, tidsstämplar på nanosekunden, symlänkar och katalogposter — se bred-provning, avsnitt 5.

- Arkiven i protokollet är temporära; deras md5 i `protokoll.json` finns

för spårbarhet men arkivstorleken är inte deterministisk mellan körningar (se bred-provning, avsnitt 2.2).