

Linux-puzzel opgelost

Johan Swenker

Welke Debian pakketten zijn later handmatig geïnstalleerd?
Een oefening met synaptic, sed en reguliere expressies.

Op een bijeenkomst van de HCC in Assen vroeg iemand: 'Voordat ik Linux opnieuw installeer, zou ik eigenlijk wel een lijst willen hebben van alle software die ik later toegevoegd heb. Is zoiets mogelijk?'

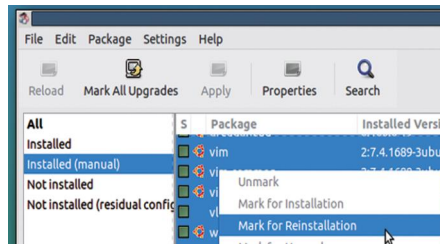
Het was even puzzelen, maar we zijn eruit gekomen. Althans, voor een voldoende moderne distributie die op Debian gebaseerd is, zoals Ubuntu 16.04 en Raspbian GNU/Linux 8.

Met het commando `dpkg` kun je Debian pakketten manipuleren. Je kunt er ook de status van een pakket mee opvragen. Met `'dpkg -l'` krijg je een lijst van alle pakketten die geïnstalleerd zijn. Maar helaas kun je met `dpkg` niet de vraag beantwoorden of het pakket automatisch geïnstalleerd is bij het installeren van Linux, of dat je het later toegevoegd hebt.

Als je met `'apt-get install pakket'` wel eens een pakket toegevoegd hebt, dan heb je vast wel gezien dat `apt-get` meldt dat het pakket nu de status: *'handmatig geïnstalleerd'* heeft gekregen. Blijkbaar moeten we een commando uit de `apt`-reeks van commando's gebruiken. Inderdaad `'apt show bash'` meldt onder andere: *'APT-Manual-Installed: yes'*. Nu heb ik `bash` niet handmatig geïnstalleerd, maar het is wel de informatie die `apt` me geeft. En daar moet ik het dus mee doen.

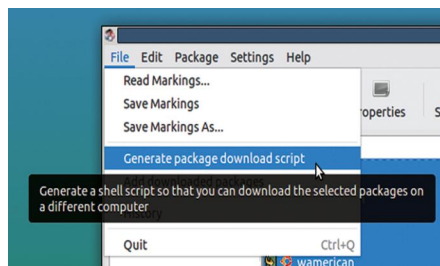
Toen we deze informatie bij elkaar verzameld hadden, wisten we: het gaat lukken. In het ergste geval schrijven we een programma dat met `'dpkg -l'` de lijst van geïnstalleerde pakketten ophaalt, om daarna per pakket met `'apt show pakketnaam'` te achterhalen wat de status is. Maar, kan het eenvoudiger? Oudere versies van Ubuntu beschikten over Synaptic, dat is een programma met een grafisch interface om pakketbeheer te doen. In Ubuntu 16.04 hebben we dit programma geïnstalleerd met `'sudo apt-get install synaptic'`. Als je het daarna als gewone gebruiker opstart, kun je geen wijzigingen aanbrengen, maar dat was ook niet de bedoeling. Maar kijk eens naar het gebruikers-interface van Synaptic: in de linker kolom kun je *'installed (manual)'* selecteren, en rechts kun je met de muis van het begin tot het eind alle pakketten selecteren, gelimiteerd tot de pakketten die, volgens het systeem, handmatig geïnstalleerd zijn. Met de

rechtmuisknop kun je al die pakketten vervolgens *selecteren voor herinstallatie*.



Wees gerust, we gaan dat niet doen. We kunnen het niet eens doen, omdat we als gewone gebruiker geen wijzigingen kunnen aanbrengen. Wat we wel kunnen doen is een *script maken* dat die wijzigingen kan aanbrengen.

Op mijn computer bestond het script uit meer dan tweehonderd regels in de vorm `'wget -c http://n1.archive.ubuntu.com/ubuntu/pool/universe/x/xosview/xosview_1.17-2_amd64.deb'`



Dit kun je met elke editor omvormen tot enkel de pakketnaam `xosview`. Maar om zelf met de hand al die duizenden lettertjes weg te halen, en er niet één te veel weg te halen, dat is vragen om problemen. Bovendien: we hebben die computer niet voor niets. Daarmee moet het veel makkelijker kunnen, en dat kan ook. Elk Linux-systeem beschikt over `sed`, een stream-editor. Je stuurt er een stroom data doorheen, en ondertussen wordt de data gemanipuleerd. De manipulatie die wij nodig hebben is het verwijderen van stukjes tekst. Binnen `sed` doe je dat door een vervang-opdracht, waarbij je de geselecteerde tekst vervangt door niets.

Met `'cat script_door_syntaptic'` krijgen we de stroom data die door `sed` verwerkt kan worden. Laten we eerst alles wissen vanaf de underscore `'_'` tot aan het einde van de regel. Daar moeten we een zogenaamde reguliere expressie voor schrijven. Eigenlijk ken

je reguliere expressies al een beetje: als je sterretjes gebruikt in `*.jpg` om al je plaatjes aan te duiden, dan heb je al een reguliere expressie gebruikt.

`Sed` gebruikt iets vergelijkbaars, maar helaas niet precies hetzelfde. Met de reguliere expressie `'_.*$'` geef je alles aan vanaf de eerste underscore (die staat er letterlijk als een `'_'`) tot aan het eind van de regel (dat eind van de regel wordt aangegeven met het dollarteken). De losse punt past bij elke letter. En het sterretje betekent zo ongeveer 'zo vaak als nodig is'. Het hele `sed` commando wordt nu:

```
sed -e s/_.*$//
```

Dit zal de tekst die tussen de eerste twee schuine strepen staat, vervangen door de tekst die tussen de laatste twee schuine strepen staat. Inderdaad, daar staat niets tussen.

Op een soortgelijke manier verwijderen we ook alle tekst tot aan de laatste schuine streep. Daar hebben we de reguliere expressie `'^.*$'` voor nodig. Waar we eerst het dollarteken gebruikten om tot aan het eind van de regel te gaan, gebruiken we nu het 'dakje' om vanaf het begin van de regel te selecteren.

We hebben nu echter een extra probleem: de schuine streep die we zoeken, die hebben we ook gebruikt om de zoekstring en de vervangstring mee te markeren. De schuine streep als markerings-teken is slechts conventie. Iedereen gebruikt het, tot het onhandig wordt. Je mag elke letter als markerings-teken gebruiken. Het `sed`-commando om het begin te wissen kan er bijvoorbeeld als volgt uit zien: `'sed -e sX^.*$//'`. Inderdaad, ik heb de hoofdletter `X` gebruikt als markerings-teken.

Nu hoeven we alleen nog maar alles achter elkaar te plakken: `'cat script_door_syntaptic | sed -e s/_.*$// | sed -e sX^.*$// > lijst'` en we hebben een lijst met alle pakketten die handmatig geïnstalleerd zouden zijn.

Let er op dat je die lijst naar een andere computer kopieert voordat je met de nieuwe installatie begint. Bij die nieuwe installatie kun je met `'sudo apt-get install 'cat lijst''` in een klap al die pakketten opnieuw installeren.