

Accurate Backup - sposób na dokładną kopię bezpieczeństwa

2009-07-11 autor: **gani**

Jedną z nowych funkcjonalności wersji Baculi 3.0.0 jest **Accurate Backup**. Udostępnia on możliwość rozszerzonego śledzenia zmian po stronie klienta takich, jak zapamiętanie informacji o plikach usuniętych czy przeniesionych do innej lokalizacji. Nie jest to nowy typ kopii bezpieczeństwa, ale rozszerzenie możliwości kopii przyrostowych i różnicowych.

Często spotykana procedura do określania archiwizowanych danych przez program backupowy w przyrostowych bądź różnicowych kopiach bezpieczeństwa oparta jest na porównywaniu czasów stworzenia i modyfikacji pliku z czasem wykonania ostatniej pełnej kopii bezpieczeństwa. Porównywane są czasy CTIME (czas stworzenia pliku) i MTIME (czas modyfikacji pliku). Jeśli dla pliku jeden z tych dwóch czasów jest późniejszy od czasu ostatniej poprawnie wykonanej kopii bezpieczeństwa, wtedy będzie on brany pod uwagę w procesie archiwizacji. Ten sposób wykorzystywany jest również w programie Bacula. Umożliwia on archiwizację plików nowo stworzonych oraz zmodyfikowanych bez możliwości śledzenia i zapamiętania innych zmian w systemie plików klienta, jak na przykład informacji o plikach usuniętych czy przeniesionych do innej lokalizacji.

Proszę wyobrazić sobie sytuację, w której następuje utrata wszystkich danych z komputera obsługiwanego przez jednego z klientów backupowych. Administrator otrzymuje zlecenie przywrócenia utraconych danych. Wybiera on odpowiednie kopie zapasowe zawierające najaktualniejsze dane. W przypadku gdy tuż przed awarią została wykonana kopia pełna danych klienta, przywrócenie dokładnego i aktualnego stanu plików nie powinno stanowić problemu. Jednak w sytuacji, gdy od czasu wykonania ostatniej kopii pełnej wykonało się kilka kopii przyrostowych czy różnicowych, proces przywracania danych z odpowiednich kopii bezpieczeństwa wiąże się z przywróceniem wszystkich danych wraz z plikami, które od czasu wykonania kopii pełnej zostały celowo skasowane czy przeniesione. Będzie to więc „nadmiarowe” przywrócenie danych nie odzwierciedlające faktycznego stanu systemu plików klienta.

Jedną z nowych funkcjonalności wersji Baculi 3.0.0 jest **Accurate Backup**. Udostępnia on możliwość rozszerzonego śledzenia zmian po stronie klienta takich, jak zapamiętanie informacji o plikach usuniętych czy przeniesionych do innej lokalizacji. Nie jest to nowy typ kopii bezpieczeństwa, ale rozszerzenie możliwości kopii przyrostowych i różnicowych. Gdyby w opisanej w poprzednim akapicie sytuacji użyto archiwizacji z opcją **Accurate**, wtedy odtworzone dane byłyby dokładnym stanem plików klienta, bez nadmiarowych plików.

Opcję dokładnego backupu definiuje się dla zasobu Job poprzez dyrektywę **Accurate = yes/no**. Domyślnie dyrektywa ta ustawiona jest na **'no'**. Jeśli dyrektywa jest włączona dla zasobu Job (**Accurate = yes**), wtedy kopie bezpieczeństwa wykonane z jego użyciem będą wykonywały dokładne backupy z użyciem mechanizmu Accurate. Dla kopii pełnych (ang. full backup) dyrektywa ta nie ma znaczenia. Dla kopii przyrostowych i różnicowych przy włączonej dyrektywie, Daemon Director przy każdej archiwizacji będzie wysyłał do klienta (File Daemon) listę poprzednio zarchiwizowanych danych, której użyje klient do określenia, jakie pliki zostały dodane, jakie zostały przeniesione a jakie usunięte. To umożliwi Baculi stworzenie dokładnego backupu na systemie plików klienta w danym punkcie czasu. Jeśli po wykonaniu takich kopii bezpieczeństwa nastąpi przywrócenie danych, dane przywrócone będą odzwierciedlały dokładny stan plików z danego punktu czasu.

Poniższy rysunek przedstawia zasadę działania **Accurate Backup**.

DZIEŃ ->	PON	WTOREK	ŚRODA	CZWARTEK	PIĄTEK
TYGODNIA					
	ala.txt	ala.txt	ala.txt	ala.txt	janek.txt
STAN ->	ma.txt	ma.txt	szarego.txt	woli.txt	woli.txt
PLIKÓW	kota.txt	szarego.txt kota.txt	kota.txt	psy.txt	zupe.txt
TYP ->	PEŁNA	PRZYROSTOWA	PRZYROSTOWA	PRZYROSTOWA	PRZYROSTOWA
KOPII					
	+ala.txt	+szarego.txt	-ma.txt	-szarego.txt	-ala.txt
PLIKI ->	+ma.txt			-kota.txt	-psy.txt
BACKUPU	+kota.txt			+woli.txt	+janek.txt
				+psy.txt	+zupe.txt

W przykładzie zakłada się, że dyrektywa **Accurate** dla zasobu Job ustawiona jest na **'yes'**. Przed archiwizacją istnieją trzy pliki `ala.txt`, `ma.txt`, `kota.txt`. Trafiają one do pierwszej pełnej kopii bezpieczeństwa w poniedziałek. Następnie zostaje utworzony plik `szarego.txt`, który zostaje zarchiwizowany przy wykonaniu drugiej kopii (przyrostowej) we wtorek. W środę usunięto plik `ma.txt` i wykonano trzecią kopię (przyrostową), do której nie trafi żaden plik, jednak zostanie zapamiętana informacja o tym, że plik `ma.txt` został usunięty. W czwartek usunięto plik `szarego.txt` i `kota.txt` oraz dodano pliki `woli.txt` oraz `psy.txt`. Te informacje zostaną zapamiętane przy czwartej kopii (przyrostowej). Przed piątkową kopią zapasową usunięto plik `ala.txt` i `psy.txt` oraz dodano pliki `janek.txt` i `zupe.txt`. Te informacje również zostaną zapamiętane.

Przywrócenie plików będzie wyglądało następująco:

- przywrócenie stanu plików z wtorku – zostaną przywrócone pliki `ala.txt`, `ma.txt`, `szarego.txt` i `kota.txt`. Ta sytuacja nie różni się niczym od działania kopii przyrostowej bez opcji **Accurate**.
- przywrócenie stanu plików z środy – zostaną przywrócone pliki `ala.txt`, `szarego.txt` oraz `kota.txt`. Tutaj już widać różnicę od zasady działania standardowej kopii przyrostowej, ponieważ nie został przywrócony plik `ma.txt`. Oznacza to, że Bacula poprawnie rozpoznała, że we środę przy wykonywaniu kopii bezpieczeństwa, nie było w systemie plików pliku `ma.txt`. Uzyskany stan danych jest dokładnym stanem plików jaki był u klienta w środę.
- przywrócenie stanu plików z czwartku – zostaną przywrócone pliki `ala.txt`, `woli.txt` oraz `psy.txt`. Gdyby do archiwizacji użyto zasobu Job bez włączonej opcji **Accurate**, przywrócono by pliki: `ala.txt`, `ma.txt`, `szarego.txt`, `kota.txt`, `woli.txt`, `psy.txt`. Porównując te dwa wyniki, widać że w przywróceniu z kopii bezpieczeństwa bez ustawionej opcji **Accurate** byłyby trzy pliki nadmiarowe. Jest to dwa razy więcej plików, niż w przywróceniu plików z kopii z zaznaczoną opcją **Accurate**. To jest właśnie ta nadmiarowość, o której wspominałem na początku artykułu.
- przywrócenie stanu plików z piątku – zostaną przywrócone pliki `janek.txt`, `woli.txt` oraz `zupe.txt`, czyli dokładny stan plików jaki występował u klienta w piątek. W tym przypadku, przy wykorzystaniu kopii bez opcji **Accurate**, zostałyby przywrócone pliki: `ala.txt`, `ma.txt`, `kota.txt`, `szarego.txt`, `woli.txt`, `psy.txt`, `janek.txt`, `zupe.txt`. Jest to prawie trzy razy więcej niż w sytuacji użycia opcji **Accurate**.

Podsumowując, **Accurate Backup** daje możliwość przywrócenia dokładnego stanu danych z dowolnego czasu, zapobiega przywracaniu nadmiarowych plików. Może więc przydać się na przykład do archiwizacji całego systemu plików lub tam, gdzie archiwizowane pliki są często zmieniane przez użytkownika. Funkcjonalność ta posiada jednak pewne ograniczenia.

Przed użyciem **Accurate Backupu** trzeba wiedzieć, że wymaga on dużych zasobów systemowych (procesor i pamięć operacyjna) zarówno po stronie daemona Director jak i klienta (File Daemon). Dotyczy to głównie wygenerowania listy plików z poprzednich kopii zapasowych oraz porównywania każdego pliku z zasięgu określonego w zasobie FileSet z taką listą. Jeśli pod uwagę brane są miliony plików, potrzeba będzie dużo pamięci operacyjnej.

Drugie ostrzeżenie dotyczy archiwizowania z użyciem pluginu, który nie jest zaprojektowany do obsługi funkcjonalności **Accurate Backup**. Późniejsze przywrócenie danych z takich kopii bezpieczeństwa może przywrócić niepoprawny stan plików.

opracowano na podstawie [Concepts and Overview Guide](#)