

УДК 004.051

Методы организации хранения метаданных в современных форматах архивации данных

А. М. Неганов

Московский физико-технический институт (государственный университет)

alexey.neganov@phystech.edu

Метаданные архива – это данные о содержащихся в нём файлах и блоках данных, включающие имя и разнообразные атрибуты файла, данные, задающие внутреннюю иерархическую структуру архива, а также данные, относящиеся к логике работы архиватора.

Скорость работы с метаданными, определяемая принципами организации их хранения для данного формата, является одним из ключевых параметров, влияющих на быстродействие операций над архивом в целом, таких как создание архивной копии, её обновление, восстановление данных, просмотр и управление содержимым архива и пр.

Работа посвящена исследованию способов хранения метаданных архива и управления ими. Целью исследования было определение наиболее производительных и хорошо масштабируемых при увеличении объёмов данных и нагрузки методов, применяемых в современных открытых форматах, для чего был изучен ряд современных продуктов для архивации данных.

Литература

1. *Brandt S., Long D., Miller E., Xue L.* Efficient Metadata Management in Large Distributed Storage Systems // Proceedings of the 20th IEEE / 11th NASA Goddard Conference on Mass Storage Systems and Technologies. — 2003.
2. *Jiang H., Wang J., Xian F., Zhu Y.* HBA: distributed metadata management for large cluster-based storage systems // IEEE transactions on parallel and distributed systems. — 2008. — V. 19. — No. 4.
3. Bacula Developer's Guide. <http://www.bacula.org/7.0.x-manuals/en/developers/developers.pdf> — 2015.