

**УДК 004.6**

Кравцов Максим Дмитриевич, студент, магистрант  
(КузГТУ, г. Кемерово)  
Kravtsov Maxim D, student, master  
(KuzSTU, Kemerovo)

## **ПРОБЛЕМА ФОРМАТНОЙ ПРОВЕРКИ БОЛЬШИХ ОБЪЕМОВ ИНФОРМАЦИИ**

### **PROBLEM OF FORMATTING VERIFICATION OF LARGE VOLUMES OF INFORMATION**

Развитие электроники, компьютерных и информационных технологий за последние годы привели к тому, что появились проблемы с хранением, обработкой и форматной проверкой информации. Для хранения и обработки существует множество готовых решений, позволяющих распределять данные по кластерам так, чтобы время на ее поиск и обработку свести к минимуму, но перед тем, как сохранить или обработать информацию, необходимо убедиться в том, что ее формат соответствует его описанию. На сегодняшний день, это одно из не многих слабых мест в области обработки и хранения информации. Для ее проверки может потребоваться много времени, а может случиться так, что ЭВМ не хватит ресурсов. Встает вопрос, почему проверка занимает много времени? Ответ прост – для форматной проверки, скажем xml-документа, его необходимо полностью выгрузить в оперативную память, а как вы понимаете, это занимает какое-то время, и, если оперативной памяти не хватит, для ее выгрузки, мы увидим банальную ошибку о нехватки памяти.

Для решения такой проблемы можно найти много решений, например, увеличить объем оперативной памяти на ЭВМ, но что, если этого будет недостаточно? Основная проблема в том, что для проверки всего документа, как уже говорилось выше, его необходимо выгрузить в оперативную память и после этого, приступать к его проверке. Очевидно, что необходимы программные решения, способные обрабатывать большие объемы информации, при ограниченных ресурсах.

На сегодняшний день, на рынке имеются программно-аппаратные решения, позволяющие обрабатывать информацию при ограниченных ресурсах достаточно быстро, однако покупать их не всегда рационально, а реализация программных решений, позволяющих выполнить проверку данных при ограниченных ресурсах нет. Таким образом, появляется необходимость в программном решении, которое позволит ускорить процесс форматной проверки информации в условиях ограниченных

ресурсов ЭВМ с возможностью интеграции в готовые решения по ее обработке и хранению.