

---

Subject: Re: Substring search algorithm

Posted by [mirek](#) on Sun, 26 Feb 2012 12:14:06 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Quite interesting idea. Have to admit took me hour+ to understand the algorithm; perhaps part of initial misunderstanding was that if I understand it well, the substring length is limited to 64KB (or `H_size`) (code suggest something like `offset_t` maximum).

Also, even the the original code suggests that it only works fine for searched strings > 20kb, because initialization costs are pretty high - that is a bit hight for `Upp::String`.

OTOH, it might be interesting to try this with `VectorMap` instead of that hash thing in the code. Very likely, it would be quite faster and init costs would be much smaller.

Mirek

---