Subject: Re: Efficient pattern-matching in a large array
Posted by Foldy Lajos on Tue, 02 Aug 2011 20:04:19 GMT
View Forum Message <> Reply to Message

On Tue, 2 Aug 2011, Chris O'Dell wrote:

> I have a byte array of tens of millions values (read from a binary
> file). I want to find an 8-byte pattern where ever it occurs. Similar
> to where, but 8 values at a time instead of one. The problem is
> finding a way to do this that is as fast as possible. The pattern is:
> [234, 203, 138, 216, 21, 52, 117, 39].
>
> Right now, I have this in a big loop and it takes a while. We tried
> "match2" but it ran out of memory.
>
> Thanks,
> Chris
>
>

Hint: use ULONG64 (8 byte) comparison:

```
n=1000
barr=bindgen(n)
bpat=[42b, 43b, 44b, 45b, 46b, 47b, 48b, 49b]
lpat=fix(bpat, 0, type=15)

for j=0,7 do begin
    larr=fix(barr, j, j eq 0 ? n/8 : n/8-1, type=15)
    w=where(larr eq lpat, count)
    if count gt 0 then print, 'found:', 8*w+j
endfor
```

regards,
Lajos