
Subject: Re: looking for sort procedure
Posted by [kak](#) on Thu, 16 Jan 1997 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

"R. Bauer" <r.bauer@kfa-juelich.de> writes:

> It was surprising me that's idl's build-in sort procedure is very very
> slow.

> for this example it needs on may RS6000 AIX more than 2 minutes.

```
> a = indgen(10000)
> b = [a,a]
> print,systime(0)
> s = b(sort(b))
> print,systime(0)
```

> end

> This is much too long.

Hi all,

there seems to be a bug in the implementation of this routine:
I tested it on a SUN Ultrasparc and on an IBM RS6000, which
has about the same speed (at least for the Monte Carlo simulation
coded in F77, which usually runs on these boxes).

Result for SUN: below 1 second, 6 seconds for a=lindgen(100000L)
Result for IBM: about 145 seconds

Is this a known bug/feature? There seems to be a major problem
either in the implementation of sort or in the way it uses
the machine's resources (bad optimization?).

Karl

--

--

IPP, PO Box 1533 | Phone: +49-89-3299-1655 | E-Mail:
D-85740 Garching | FAX : +49-89-3299-1149 | krieger@ipp.mpg.de

Subject: Re: looking for sort procedure
Posted by [steinhh](#) on Thu, 16 Jan 1997 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

In article <32DD43B9.31DF@kfa-juelich.de>, "R. Bauer" <r.bauer@kfa-juelich.de> writes:
|> Hi,

```
|>
|> It was surprising me that's idl's build-in sort procedure is very very
|> slow.
|>
|> for this example it needs on my RS6000 AIX more than 2 minutes.
|>
|> a = indgen(10000)
|> b = [a,a]
|> print,systemtime(0)
|> s = b(sort(b))
|> print,systemtime(0)
|>
|> end
|>
|>
|> This is much too long.
|>
|> Did someone have a better sort routine which I can have?
```

This was a big surprise to me as well, since I've reasons to believe that IDL's sorting algorithms (at least the ones used by the MEDIAN function) are quite good, and well implemented.

One thing though - you've probably tested the routine with one of the worst input arrays available, a pathological example that illustrates that sorting algorithm performance predictions are usually statistical in nature.

Compare the the following:

```
IDL> a=indgen(10000)
IDL> b=[a,a]
IDL> t=systemtime(1) & s = b(sort(b)) & print,systemtime(1)-t
    39.870118
```

```
IDL> a=randomn(seed,10000)
IDL> b=[a,a]
IDL> t=systemtime(1) & s = b(sort(b)) & print,systemtime(1)-t
    0.13085902
```

!!!

I.e., almost 300 times faster for random data (DEC Alphaserer).

But even if your application's data looks almost like your example, all is not lost... look at this:

```
IDL> a=indgen(10000)
```

```
IDL> b=[a,a]
IDL> t=systime(1) & b=shift(b,1) & s = b(sort(b)) & print,systime(1)-t
0.64257896
```

So it shouldn't be too difficult to shake the data a bit before sorting to improve performance!

Stein Vidar H. Haugan

Subject: Re: looking for sort procedure
Posted by [Pirmin Kaufmann](#) on Thu, 16 Jan 1997 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

William Thompson wrote:

```
>
> kak@sat.ipp-garching.mpg.de (Karl Krieger) writes:
>
>> "R. Bauer" <r.bauer@kfa-juelich.de> writes:
>
>>> It was surprising me that's idl's build-in sort procedure is very very
>>> slow.
>
>>> for this example it needs on my RS6000 AIX more than 2 minutes.
>
>>> a = indgen(10000)
>>> b = [a,a]
>>> print,systime(0)
>>> s = b(sort(b))
>>> print,systime(0)
>
>>> end
>
>>> This is much too long.
>
>> Hi all,
>
>> there seems to be a bug in the implementation of this routine:
>> I tested it on a SUN Ultrasparc and on an IBM RS6000, which
>> has about the same speed (at least for the Monte Carlo simulation
>> coded in F77, which usually runs on these boxes).
>
>> Result for SUN: below 1 second, 6 seconds for a=lindgen(100000L)
>> Result for IBM: about 145 seconds
>
>> Is this a known bug/feature? There seems to be a major problem
>> either in the implementation of sort or in the way it uses
>> the machine's resources (bad optimization?).
```

>
> I also tried the above example on a DEC AXP 3000/600 where it took about 60
> seconds. I wonder if there's something in the code that is optimized for Sun
> workstations, maybe going back to the days when the first Unix port of IDL was
> called SunIDL?

Probably not. It only takes a few seconds on other non-unix, non-sun
(but, more important, non-ibm) machines.

Macintosh IIfx (40 MHz, 68030 CPU, 68882 co-processor): 5 seconds
VAXstation 4000-90A: less than 1 second

--

Pirmin Kaufmann pkaufmann@etl.noaa.gov
NOAA/ERL/ETL phone: 303 497-6613
325 Broadway R/E/ET7 fax: 303 497-6978
Boulder, CO 80303-3328

Subject: Re: looking for sort procedure
Posted by [brian.jackel](#) on Thu, 16 Jan 1997 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

In article <32DD43B9.31DF@kfa-juelich.de> "R. Bauer" <r.bauer@kfa-juelich.de> writes:

> It was surprising me that's idl's build-in sort procedure is very very
> slow.
> for this example it needs on my RS6000 AIX more than 2 minutes.

```
> a = indgen(10000)
> b = [a,a]
> print,systemtime(0)
> s = b(sort(b))
> print,systemtime(0)
> end
```

> This is much too long.

Yes, it certainly is. On a 66Mhz 486 PC running Windows 95 your example
takes less than 2 seconds. What results do you get from the time_test
benchmark?

Brian Jackel
University of Western Ontario

Subject: Re: looking for sort procedure
Posted by [thompson](#) on Thu, 16 Jan 1997 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

kak@sat.ipp-garching.mpg.de (Karl Krieger) writes:

> "R. Bauer" <r.bauer@kfa-juelich.de> writes:

>> It was surprising me that's idl's build-in sort procedure is very very
>> slow.

>> for this example it needs on my RS6000 AIX more than 2 minutes.

```
>> a = indgen(10000)
>> b = [a,a]
>> print,systemtime(0)
>> s = b(sort(b))
>> print,systemtime(0)
```

```
>> end
```

>> This is much too long.

> Hi all,

> there seems to be a bug in the implementation of this routine:
> I tested it on a SUN Ultrasparc and on an IBM RS6000, which
> has about the same speed (at least for the Monte Carlo simulation
> coded in F77, which usually runs on these boxes).

> Result for SUN: below 1 second, 6 seconds for a=lindgen(100000L)
> Result for IBM: about 145 seconds

> Is this a known bug/feature? There seems to be a major problem
> either in the implementation of sort or in the way it uses
> the machine's resources (bad optimization?).

I also tried the above example on a DEC AXP 3000/600 where it took about 60 seconds. I wonder if there's something in the code that is optimized for Sun workstations, maybe going back to the days when the first Unix port of IDL was called SunIDL?

Bill Thompson
