

# Extend checkdata to make it more useful

## Introduction

The [Pinger Data Collection](#) web page provides information on the success or otherwise of the daily data gathering process. We review this page on a regular basis and if the data has not been available from the monitor for roughly a week or more. The next step is to manually look up the email of the contact(s) at the monitoring site and email them requesting assistance and providing information on the problem (e.g. is the monitor pingable, does the ping\_data.pl web form respond etc.)

## First improvement

We need to add an icon to each line which links to creating an email to send to the contacts at the site. The contacts can be found in the NODEDETAILS database. You can access the NODEDETAILS database using [dbprac.pl](#). For example one can find the contacts for the SLAC node via:

```
http://www.slac.stanford.edu/cgi-wrap/dbprac.pl?alias=EDU.SLAC.STANFORD.N3&property=CONTACTS
```

Since the contact information is free form, after the contact information has been obtained it will be necessary to parse it to extract the valid email address (es). The module to do this should extract the userid before the @ symbol and the host name that follows the @ symbol. These should be validated to ensure the ID and name are correct. I believe the following regular expression can be used to check the hostname:

```
$hostname=~/((([a-z0-9]+)|([a-z0-9]+[-][a-z0-9]+))([.])+)/
```

Or look at <http://regexlib.com/DisplayPatterns.aspx> you may be able to find a more complete solution

The modification would be needed in [checkdata\\_gif.pl](#). It may be simpler in many cases to clean up the formatting of the email addresses in NODEDETAILS.

## Second step

Once this is done, the next step would be to automate testing so that if we have not been able to gather data from the site for the last 7 days (checkdata\_gif.pl provides the number of adjacent days we have been unable to gather data from the site) or more then see if.

1. the monitor is pingable from the host running the script,
2. whether the monitor can ping your host (traceroute.pl?function=ping), assuming it has a remote traceroute server
3. what a traceroute shows
4. the ping\_data.pl form appears, also note version of ping\_data.pl

If any of steps 1 thru 4 fail then no need to go to following steps you have already limited the scope of the problem.

Then craft an email with the information to review and send to the contacts. This email should identify what is working and where the failure is occurring. It should also point to the [FAQ](#). The email needs to come from a real person in case it gets rejected since the email To: address no longer works.

Eventually if automated the script may run on the 10th, 20th, 28th of the month. These dates are chosen to ensure there are already sufficient days in the month to be able to see enough days with no data.

## Requirements

To make the pings one can look at the existing [checkdata.pl](#) since it provides the URL to make the calls to ping and can be extended to do a traceroute. However it is probably better to make the pings from the host running the script since it reduces the number of points of failure.