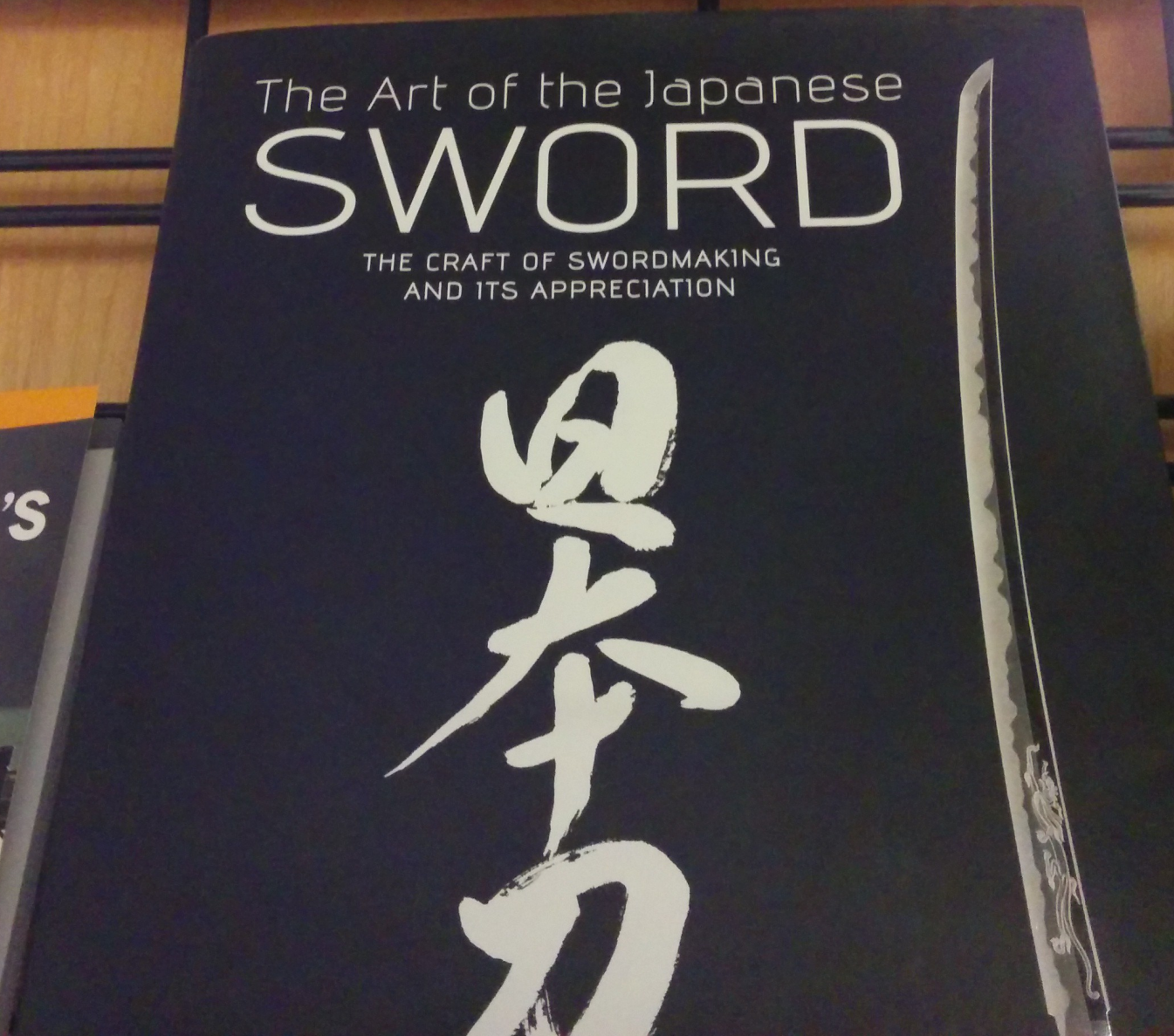
Try

find & improve some bottleneck in Debian project

Hideki Yamane
<henrich @ debian.org>

HEAVY METAL

THE
ETALWORKER'S
WORKSHOP
FOR HOME MACHINISTS


The Art of the Japanese
SWORD
THE CRAFT OF SWORDMAKING
AND ITS APPRECIATION

- There are a lot of issues among Debian project, so we need MORE contributors to fix it. But - how do we get more contributors?

- As process management (Theory of Constraints), it says "most of issues are caused by "bottleneck"". If we would fix it "bottleneck", then many issues will be gone (wow! :)

- In this session, I want to share my view/idea to find and fix bottleneck in Debian and get feedback.

$ find /debianproject -name bottleneck -print

- NEW queue
- all NEW packages should be reviewed by ftpmasters to get into repository
- ftpmasters should check whole package, then accept or reject it
- it takes many days (2,3 or 4 months, now)

wrong assumptions

- all NEW packages should be reviewed by "(only)" ftpmasters to get into repository
- ftpmasters should check "whole" package, then accept or reject it
  - if some reliable person would review package and check its review, isn't it enough?

Do NOT Burden with Yourself, please

- We are all project members,
  probably we can help you a bit :-)

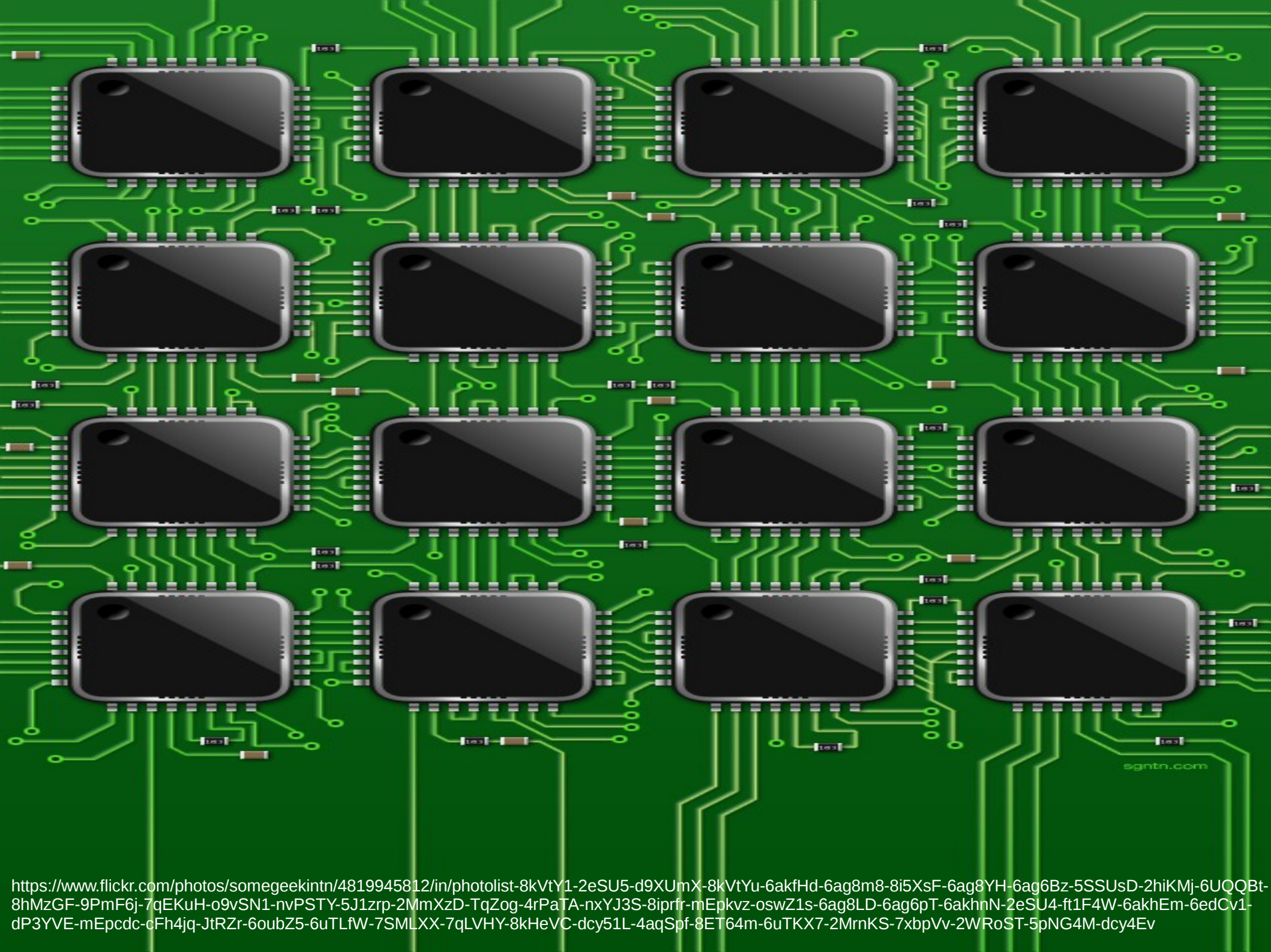
https://www.flickr.com/photos/wwarby/11494065253/in/photolist-ivG69R-9EuXjQ-caazWd-344QUT-7L2jLh-9raNuK-4NAeyW-85Zj8f-LMQZz-4UiogF-7MFnqH-HCJzd-azRhag-8t827y-bk7tHV-e4vxUD-57AGv7-gPaCyA-kPgZLc-8YzNFX-bF5gkE-auHaMM-dBUmsW-wyPhY-7a3oAE-9rdkib-dYzoJt-ndbNVx-gkMLXV-69QRUc-hLXH8U-7YDJem-nFoyGZ-akMuUU-eAFiTk-8CVMqA-8Wj4mi-6zmW1L-737YzM-7YAvRM-8Hwi9s-4nL9XH-7YAvMv-3fa8uE-7ntiiQ-7YDGKh-9sk6t4-bjxsyY-2R97kJ-6sUvAX

add "preprocessor" for NEW queue

- review by contributors
  -> can reject obvious fault earlier
- Improve
  "upload -> check -> reject -> fix -> reupload" cycle
  - don't need to wait for 2 or 3 months to be rejected ;)

from serial to parallel

- "serial" process by ftpmasters (several)
  -> "parallel" process by contributors (hundreds)
- faster (we're multi-core monster machine! :)
- less load average (pressure) for each worker



https://www.flickr.com/photos/somegeekintn/4819945812/in/photolist-8kVtY1-2eSU5-d9XUmX-8kVtYu-6akfHd-6ag8m8-8i5XsF-6ag8YH-6ag6Bz-5SSUsD-2hiKMj-6UQQBt-8hMzGF-9PmF6j-7qEKuH-o9vSN1-nvPSTY-5J1zrp-2MmXzD-TqZog-4rPaTA-nxYJ3S-8iprfr-mEpkvz-oswZ1s-6ag8LD-6ag6pT-6akhnN-2eSU4-ft1F4W-6akhEm-6edCv1-dP3YVE-mEpcdc-cFh4jq-JtRZr-6oubZ5-6uTLfW-7SMLXX-7qLVHY-8kHeVC-dcy51L-4aqSpf-8ET64m-6uTKX7-2MrnKS-7xbpVv-2WRoST-5pNG4M-dcy4Ev

contributors = buffer cache

- stable output
  - review process speed would NOT depend on ftpmasters workload by adding buffer cache

what ftpmasters should do?

- collect reviews from contributors and say "Go/Nogo(need-more-review)" to uploaders and contributors who reviewed it
  - maybe mailing list is better to share "howto" knowledge and make teach some patterns for review

reason why it's restricted to ftpmasters?

- Guess:
  probably because of "distribute" risk
  - copyright violation, patent, etc...
  - then use GPG encryption
    - first, for only @debian.org address is in keyring is ok
      – it's already "trusted" person. then expand to others...

If we do it, what is expected to happen?

- Make unstable more attractive for developers
  - All your software are belong to us :)

review contributor

- if they would be well trained, maybe would be future ftpmasters/assistants (more improvements :)

how to check it's success (or not)

- Statics:
  daily package numbers and
  average days for stay in "NEW" queue

love :-)