distcc -- A Distributed C/C++ Compiler

Martin Pool
HP

http://distcc.samba.org/

Motivation

Building large trees like Samba, GARNOME, Linux

One or two spare machines

Don’t care enough to install clustering software

Inspired by ccache

Tridge said it wouldn’t work :-)

Existing systems

Push the command across the network

Require identical machines

- Common NFS filesystem?
- Single system image?
- Same headers, library, ...

Too much hassle

Desires

Want to run jobs in parallel (using GNU Make)

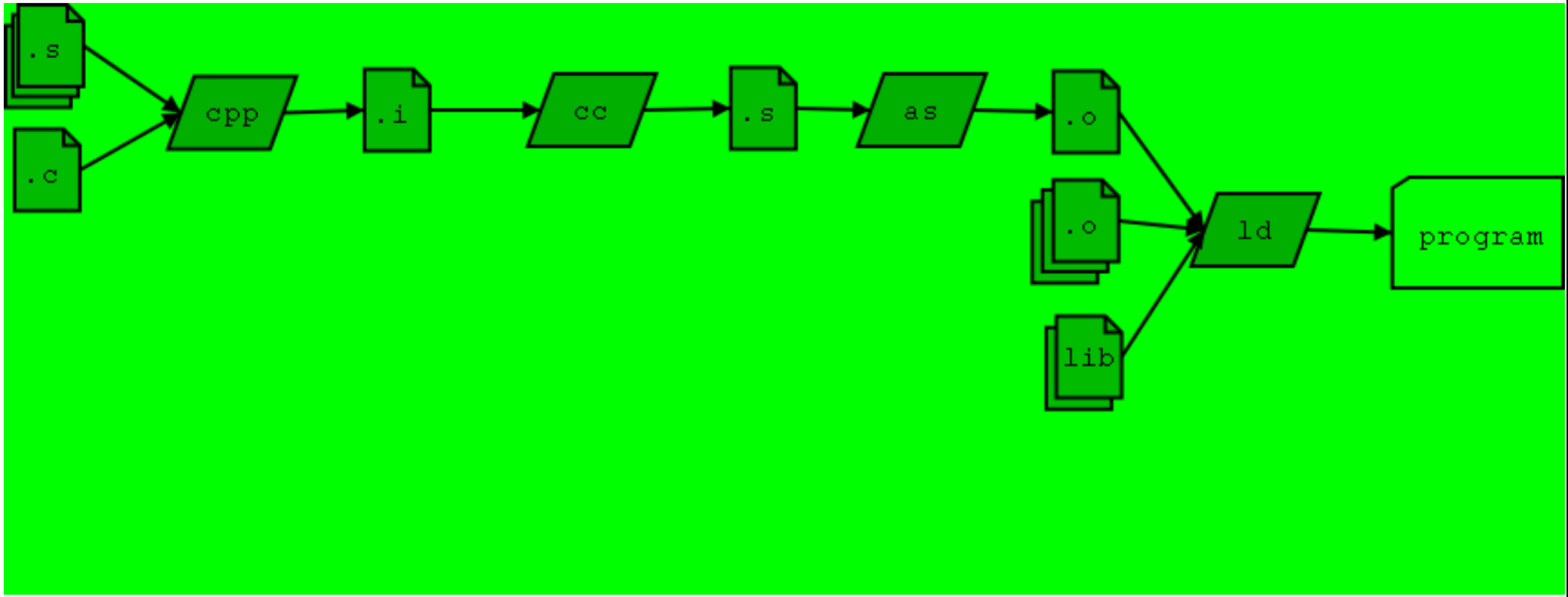
Same output

Accurate error messages

Few hassles

Assurance of correctness: simplicity, testability, ...

Phases of compilation...



Compile & assemble phases depend only on the preprocessed input, and produce a single .o output file

These tasks are "self contained", and could be farmed out

Implementation

Examine command line

Run cpp

Pick a server (least-busy-ish)

Request: preprocessed source, arguments

Response: exit code, messages, object file (if you’re lucky)

Dump back out on client

Use

Install and start daemons

export DISTCC_HOSTS=’red green blue localhost’

make CC=distcc -j6

1.0 release

Debian Sid, Mandrake, Gentoo, FreeBSD, Solaris, Zaurus...

Works across OSs, distros, and architectures

On 3 machines, 2.2 to 2.9 times faster, depending

Builds most projects: GNOME, Linux, Samba, KDE, gdb, glibc...

Caveats

C++ ABI change

Currently just plain TCP

Makefile must be reasonable ($CC, dependencies, ...)

Inefficiencies

Non-parallelizable tasks (linking, non-C stuff, ...)

Scheduler leaving machines idle

Network overhead?

CPU overhead for distcc

Test suite

2 suites: correctness, and performance

Get testability in from the word go

How to portably test an assembler?

C programs that expose knotty functions

Testing distributed & concurrent systems can be hard

Encourage useful bug reports

Future work

Run over SSH

Compression

Hide gcc warts

Better scheduling algorithm

Discovery of available machines?

More tests & benchmarking

Any questions?

http://distcc.samba.org/

mbp@samba.org

distcc@lists.samba.org

Choose Freedom