

Red Hat GFS

Richard Keech

Chief Technology Architect
rkeech@redhat.com

GFS: outline

- GFS is a clustering file system
- Unbundled product for use with Enterprise Linux
- From Sistina acquisition, 2003

based in Minneapolis, MN

founder: Matt O'Keefe

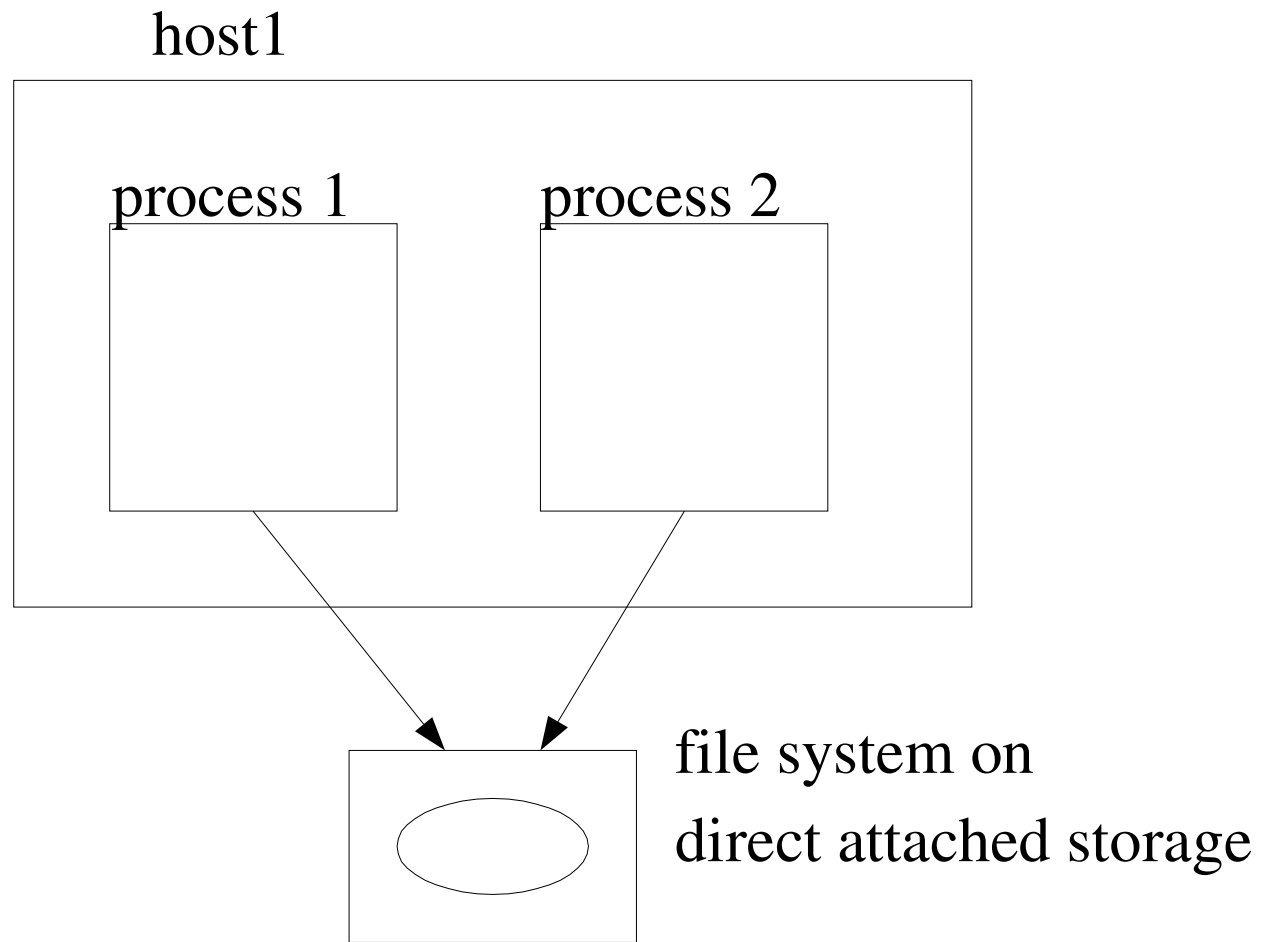
(formerly Engineering Prof at UMN)

- GFS: open source -> proprietary -> open source again.
- Now integrated with part of Cluster Suite (EL4).

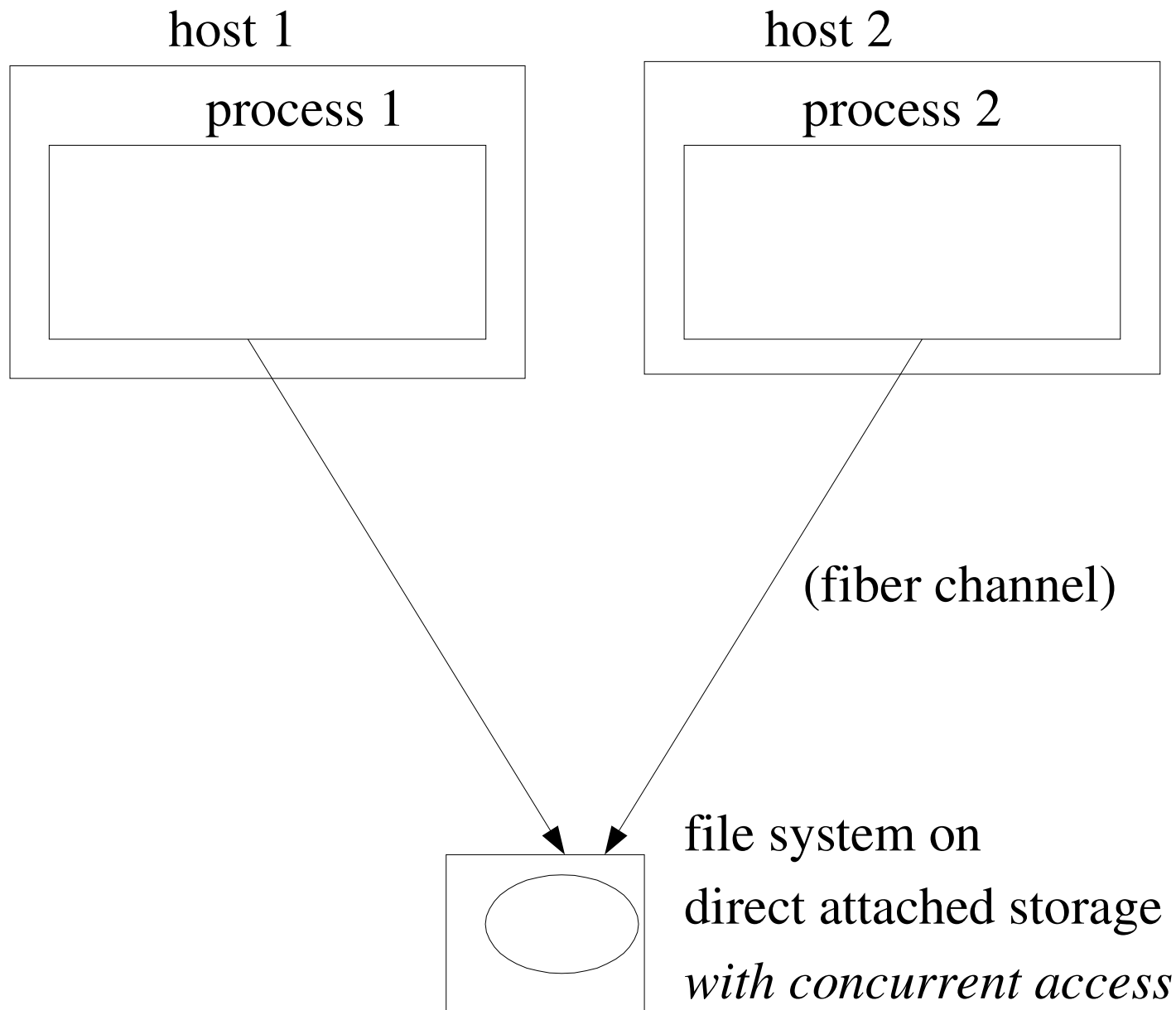
GFS: versions

- GFS version 5.x pre-Red Hat
- GFS version 6.0 for Enterprise Linux 3
- GFS version 6.1 for Enterprise Linux 4
- for use with Enterprise Linux AS, ES, WS

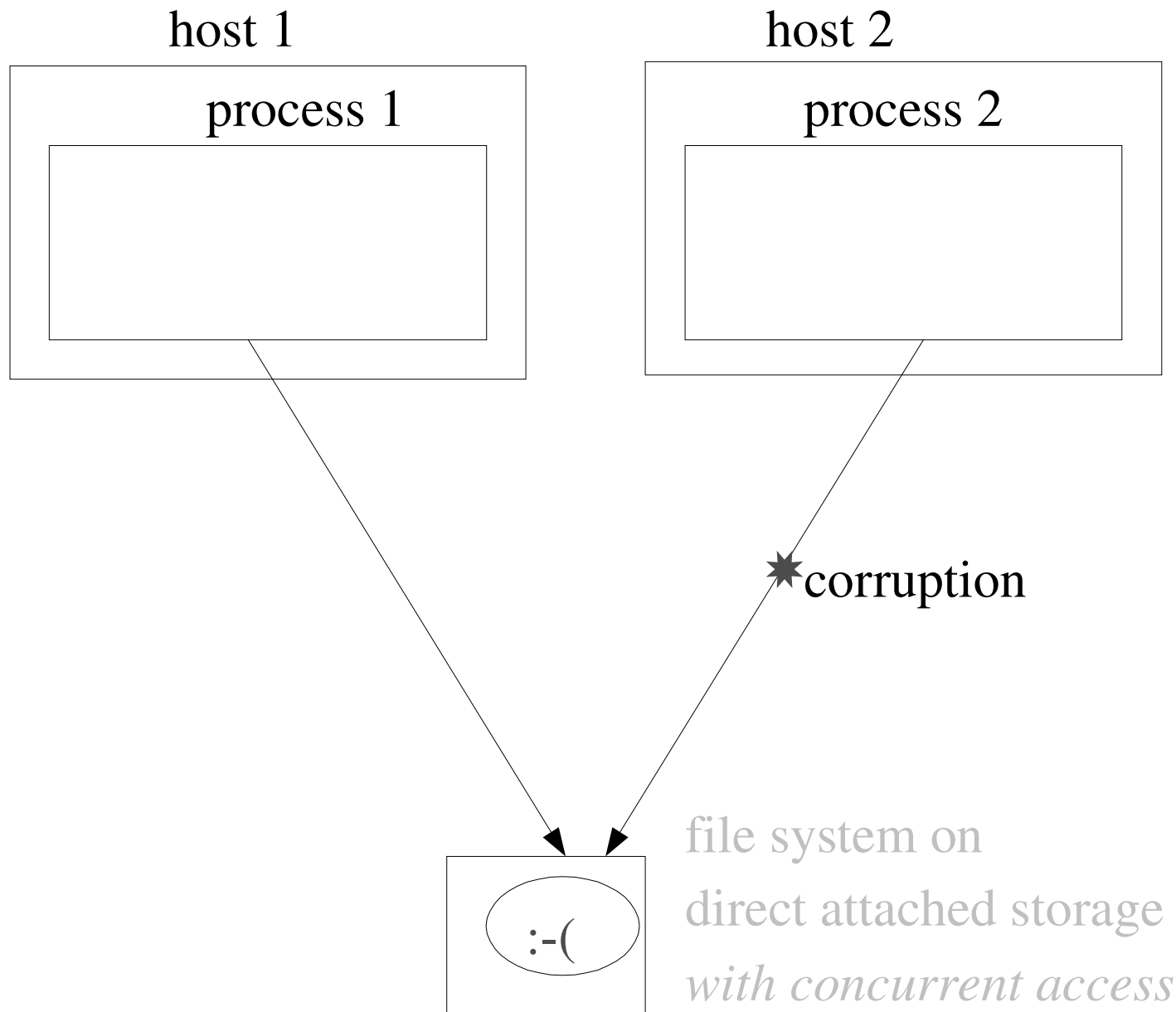
Conventional file system access



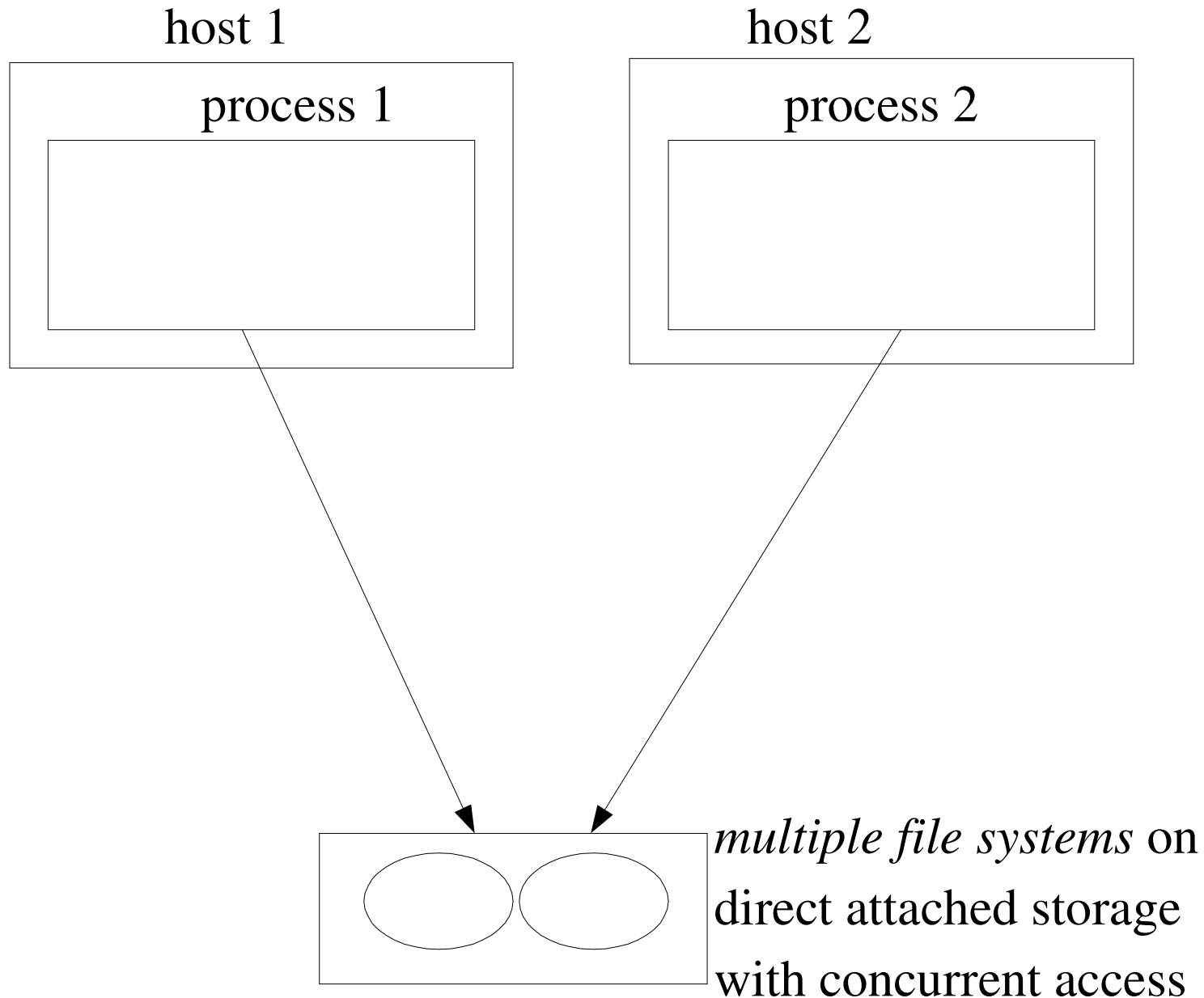
Scale out problem



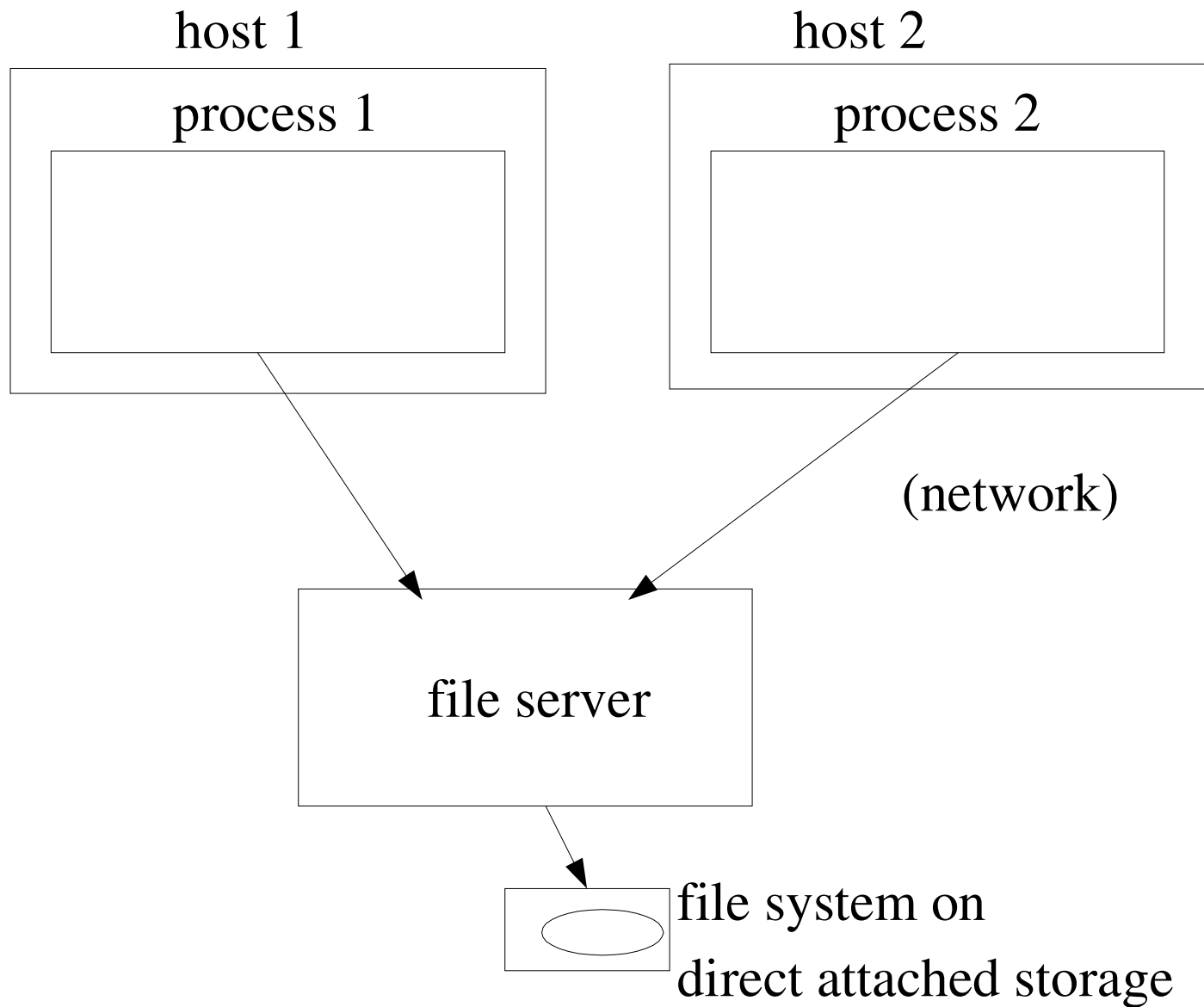
Scale out problem



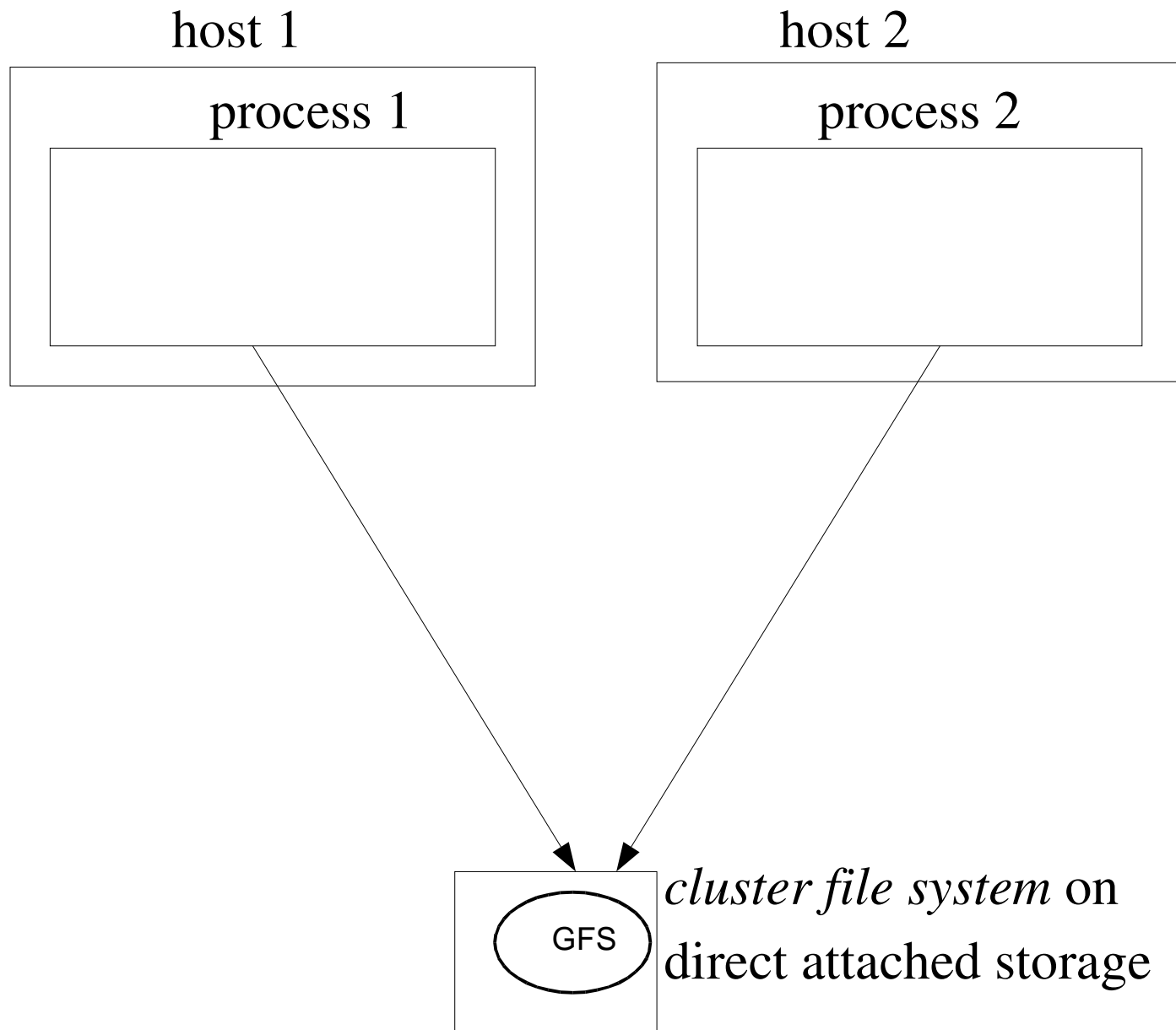
Managing the scale out problem (1)



Managing scale out (2)



Managing scale out with GFS



GFS: capability and benefit

- Many hosts can share files on same disk
 - no file server host required
- Architectural simplification
- File locking problems reduced
- File caching and coherency problems reduced
- Remove file server as a point of failure

GFS: file system characteristics

- Journalling
- Capacity to grow while live and mounted
- Context-dependent path names
- Scaling very large
- POSIX ACLs

GFS: supported architectures

- Enterprise Linux 3: x86
- Enterprise Linux 4: x86, x86_64, ia64

GFS: configurations

- minimum three nodes for quorum
- scales to hundreds of nodes
- locking nodes and storage nodes can be same or separate
- additional nodes can be added
- hundreds of nodes supported

GFS: volume manager

- “pool” simple, cluster-aware, volume manager
- pool device eg, */dev/pool/pool0*
- supports concatenation and striping
- pools are growable
- pool, software raid and LVM being merged into LVM2 and device mapper (DM) in EL4.

GFS: Storage environment

- Fiber channel SAN storage is the usual way.
- iSCSI now supported (since EL3 Update 4)

GFS: Lock management

- GFS 6.0 has redundant lock management
 - single active lock manager host
 - no single point of failure
- GFS 6.1 has distributed lock management
 - no bottleneck on lock-management host
 - shared infrastructure with cluster manager
- Have separate file and lock traffic for best performance
 - dedicated locking network segment

GFS: Fencing

- Fencing is the act of deliberate exclusion (fencing off) of uncontactable nodes.
- Fencing in GFS is not the STONITH style in RHCM.
- Fencing allows the safe recovery of locks associated with a failed node.
- Protecting against “split brain” situations is a key precaution; quorum applies.
- Methods: Network Power Switch, SAN switch.

GFS: Context dependent paths

- Imagine /data is shared by host1 and host2.
- Directory /data/host/ is to be host-specific for both.
- ```
host1# mkdir /data/host1
```

```
host2# mkdir /data/host2
```

```
host1# ln -s /data/@hostname /data/host
```
- host1 will see */data/host* as being */data/host1/*
- host 2 will see */data/host* as being */data/host2/*