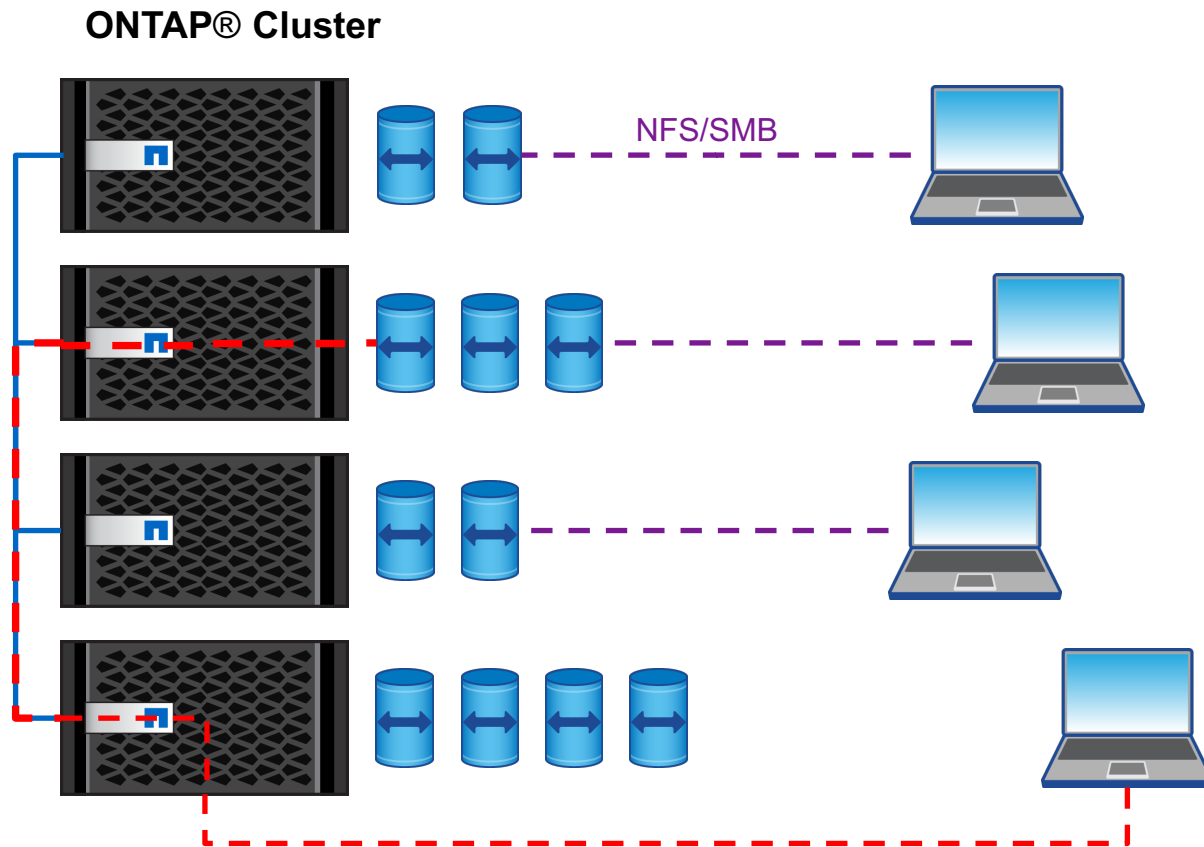# FlexGroup Volumes

A Distributed WAFL File System

Ram Kesavan, Jason Hennessy, Richard Jernigan,
Peter Macko, Keith A. Smith, Daniel Tennant, Bharadwaj V. R.
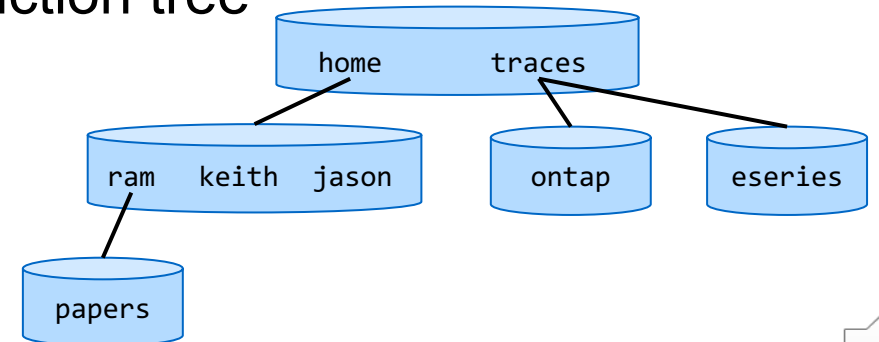
USENIX ATC '19

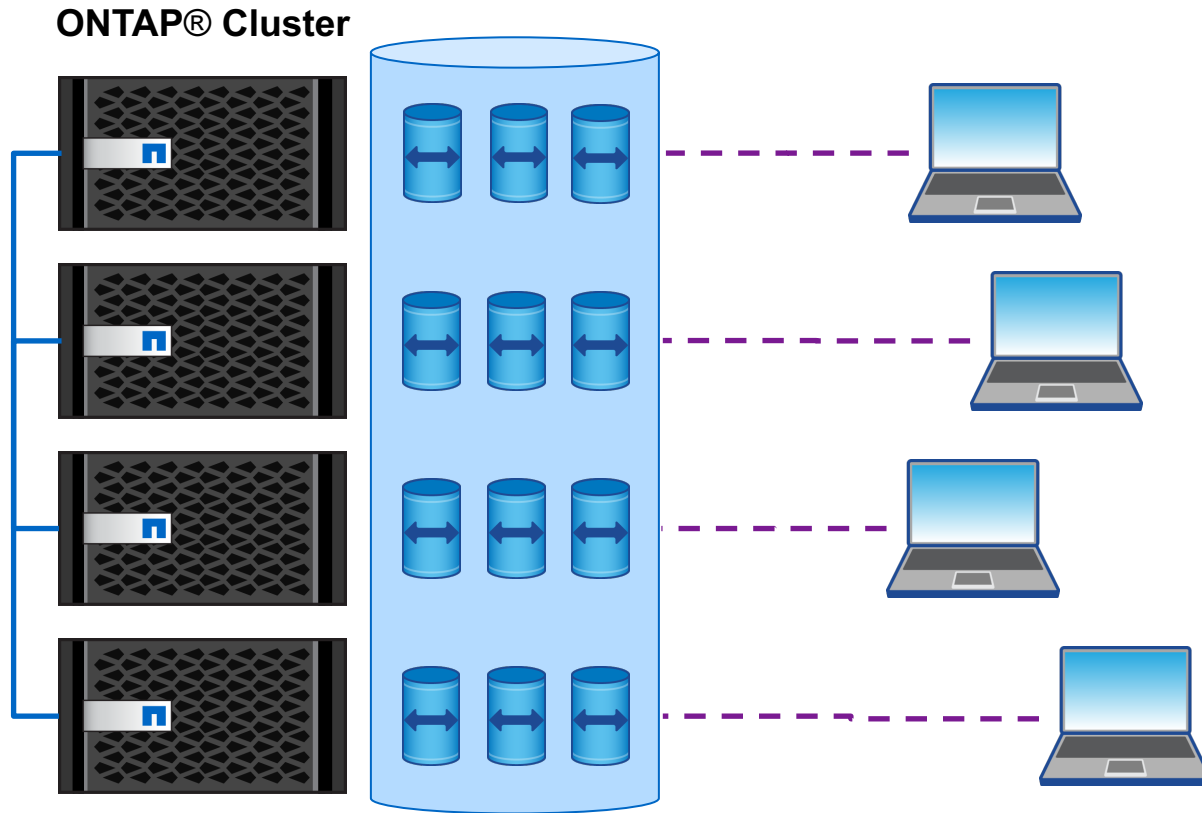# Where we were

Individual WAFL volumes

- Individual WAFL volumes managed by a single controller
- Clients mount individual volumes
- If a client connects to a different node, requests will be forwarded over the cluster interconnect
- Can join multiple volumes into a junction tree

FlexGroup® Volumes

A distributed file system

ONTAP® Cluster

- FlexGroup® = a distributed file system internally composed of WAFL volumes

- FlexGroup presents the member volumes as a single volume to clients

- ONTAP *automatically* manages file placement to spread the load

NetApp

FlexGroup® Volumes

A very high-level view
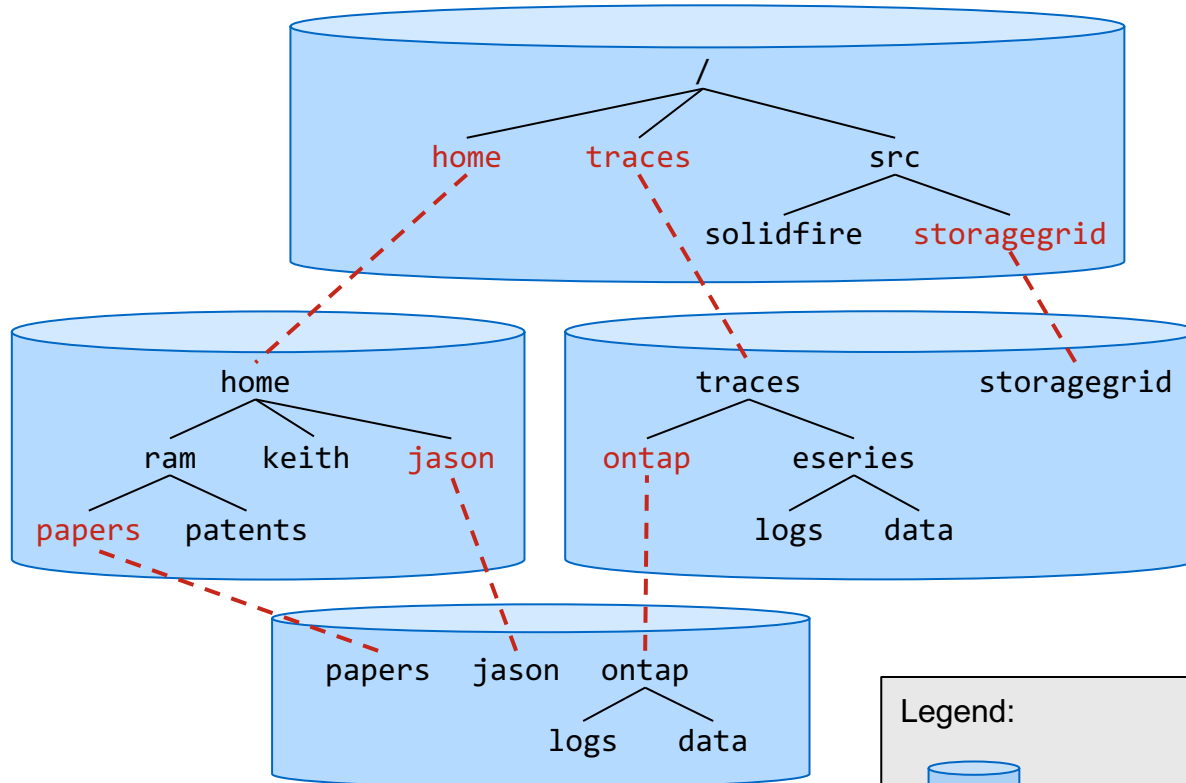
- No central metadata server

- Files and directories are stored in their entirety on the member volumes

- Remote links are *automatically* created to connect subtrees on the individual volumes into a single namespace

FlexGroup® Volumes

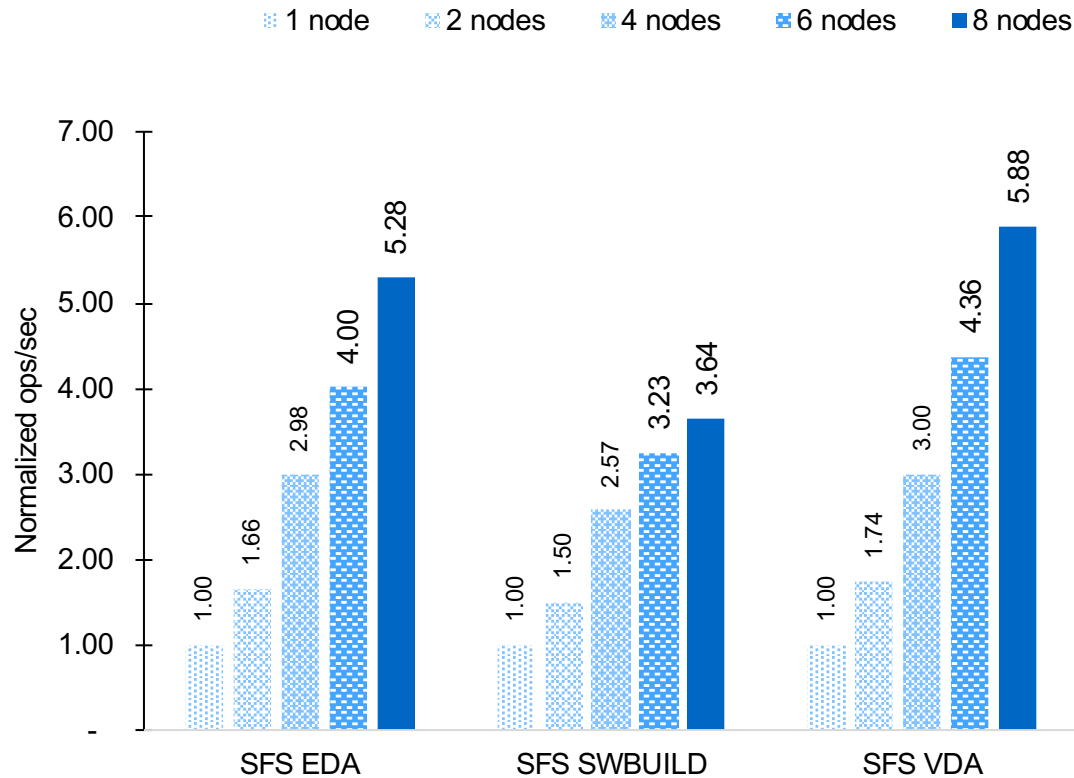Two key ingredients

Remote Access Layer

Atomically create, update, and delete remote links across volumes.

Data Placement Heuristics

Determine data placement, balancing data locality and spreading load and space utilization across the cluster.

Have we been successful?

Evaluation sneak peak

Legend: 1 node, 2 nodes, 4 nodes, 6 nodes, 8 nodes

Chart (Normalized ops/sec):

SFS EDA: 1.00, 1.66, 2.98, 4.00, 5.28
SFS SWBUILD: 1.00, 1.50, 2.57, 3.23, 3.64
SFS VDA: 1.00, 1.74, 3.00, 4.36, 5.88

- Scalable file system performance

- Thousands of FlexGroup volumes deployed storing hundreds of petabytes

- Customer data show effectiveness of the data placement heuristics

■ NetApp

NetApp

Would you like to know more?

Come to the talk:
Wednesday, July 10th, 2:20pm
ATC '19, Track I (Filesystems)

Read the paper:
https://www.usenix.org/conference/atc19/presentation/kesavan