



STORAGE DEVELOPER CONFERENCE

SNIA ■ SANTA CLARA, 2015

Instantly Finding a Needle of Data in a Haystack of Large-Scale NFS Environment

Gregory Touretsky

INFINIDAT



Who am I?

- ❑ 1997-2015: IT Systems Engineer / Solutions Architect at a Fortune 100 company
- ❑ As of 9/20: Product Manager at Infinidat

Introducing INFINIDAT

- ❑ Founded in 2011
- ❑ >100 patents
- ❑ Production deployments at Fortune 500 companies
- ❑ Over 200PB deployed in the field
- ❑ InfiniBox:
 - ❑ 99.99999% uptime
 - ❑ 2PB in 42U rack
 - ❑ Multi-protocol
 - ❑ 8kW max power consumption
 - ❑ 750k IOPS
 - ❑ Zero-wait RESTful API



INFINIDAT NAS

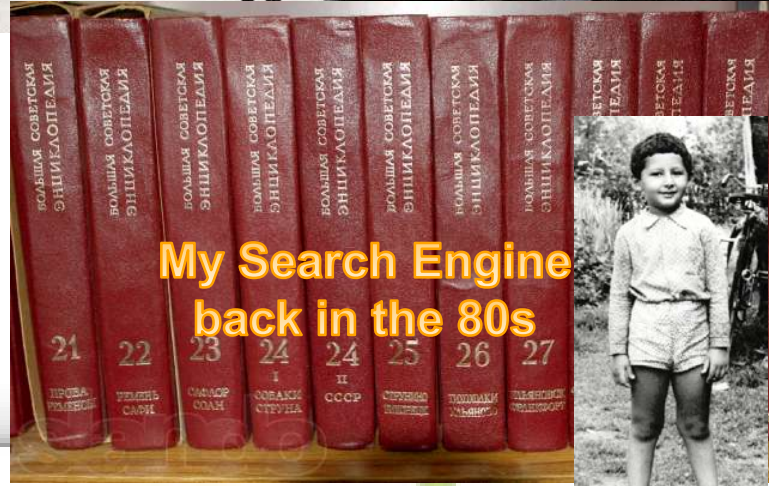
- ❑ >250,000 Spec SFS ops/sec at first release
- ❑ Infinidat Hyper-scale filesystem
 - ❑ Scales to billions of files per FS
 - ❑ 4,000 → 100,000 max filesystems
 - ❑ Virtually unlimited size
- ❑ N+2 architecture
- ❑ InfiniSnaps
 - ❑ No performance impact

And now to my past life experience...

Looking for answers?

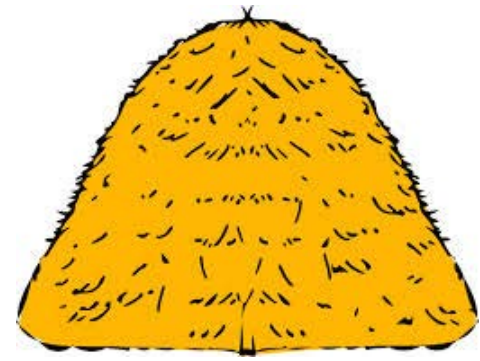
EVERY FAMILY with **CHILDREN** in it
should own and
can own
this Newest, Greatest
**ENCYCLOPAEDIA
BRITANNICA**

EVERY FAMILY, and
above all, every family
with children in it,
should own the great new
Encyclopaedia Britannica—the
one essential book for the home
—the one work bringing to
young and old the limitless
advantages of *modern knowl-
edge.*



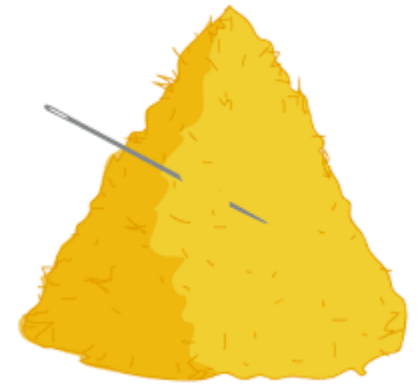
**My Search Engine
back in the 80s**

How big is the Haystack?



- ❑ Over 1,000 NFS file servers
- ❑ Over 35PB of configured capacity
- ❑ 125,000 NFS file systems
- ❑ Over 90,000 NFS clients
- ❑ 103B files in the Unified Name Space

What are the Needles?



- ❑ Comments within code
- ❑ Keywords
- ❑ Where is environment variable defined?
- ❑ Errors in logs
- ❑ Finding all references to a variable, method, or class throughout a project

`grep -R <text> <path>`



Is there a better way?

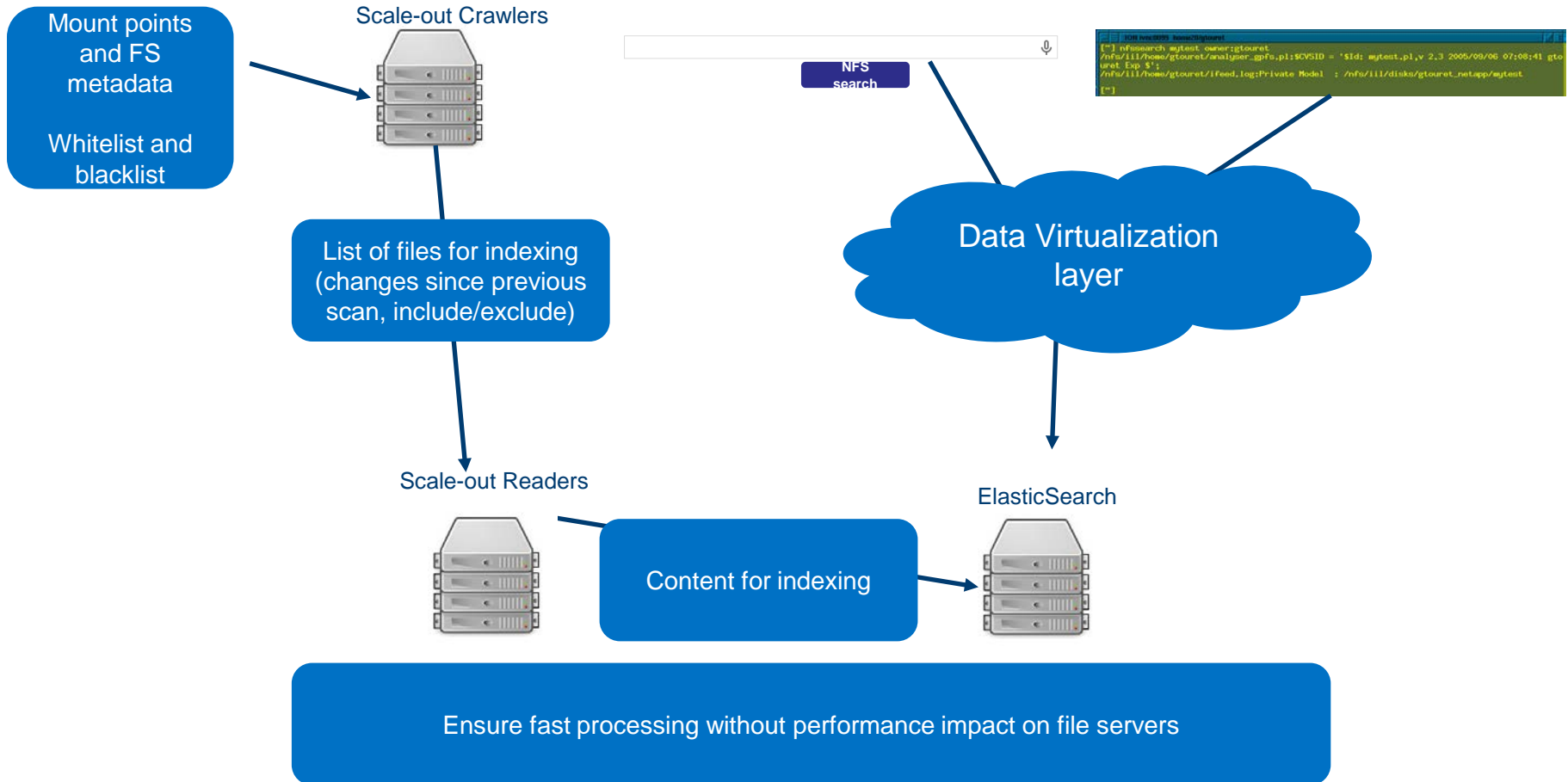
Requirements and basic assumptions

- ❑ NAS appliance = “black box”
- ❑ Indexing SLA = 24 hours
- ❑ Avoid NAS DoS due to indexing
 - ❑ Differentiate scanning load by NAS type
- ❑ Reuse
- ❑ Search “hints”: owner, time, project, ...
- ❑ Future: one interface to search through all data sources (SharePoint, Wiki, Web, DB, NFS, ...)

Considered solutions

- ❑ Enterprise search (Google Appliance)
 - ❑ Cost
 - ❑ NFS integration
- ❑ Open source – Solr
 - ❑ Concerns with scalability, supportability
 - ❑ Less applicable for other potential use cases (logs, etc)
 - ❑ Missing scalable crawler
- ❑ Open source – Elasticsearch
 - ❑ Missing scalable crawler

Implementation architecture



Pools of crawlers, readers

- ❑ One job per file system
 - ❑ Multi-threaded crawler
- ❑ Internal batch queues management system
 - ❑ Queue per file server
 - ❑ “MaxRunning” limit per file server
 - ❑ Dynamic vs. static limit
- ❑ Root access to NFS

Indexing considerations

- ❑ One index per filesystem
- ❑ White list, black list support
- ❑ Index updates
- ❑ Elasticsearch cluster configuration

User interface

- ❑ WebUI
- ❑ CLI
 - ❑ Grep-like for backward compatibility
 - ❑ Expose full Lucene interface
 - ❑ title:"The Right Way" AND text:go
 - ❑ "jakarta apache"~10
 - ❑ ...

Security aspects

- ❑ Index encryption
- ❑ ElasticSearch cluster access control
- ❑ Search access control – preserve NFS permissions
 - ❑ Initially – at the index (filesystem) granularity level

Pilot status

- ❑ One project's data is indexed
 - ❑ 85 file systems / 71TB / 409M files
- ❑ Index size: 17TB
- ❑ Access control
- ❑ UI

It works!

Before:

```
grep -R myKeyWord /nfs/mymount
```

```
<results ... results ... results>
```

32:41.11

After:

```
ESgrep myKeyWord /nfs/mymount
```

```
<results ... results ... results>
```

0:00.49

```
ESgrep myKeyWord
```

```
<results ... results ... results>
```

0:00.52

17

What if NAS is not a “black box”?

Find changes within the filesystem

- ❑ Indexing
- ❑ Data replication (rsync, etc)
- ❑ Data aging / recycling
- ❑ Controlled technology
- ❑ Antivirus
- ❑ Backup
- ❑ ...



Find changes within the filesystem

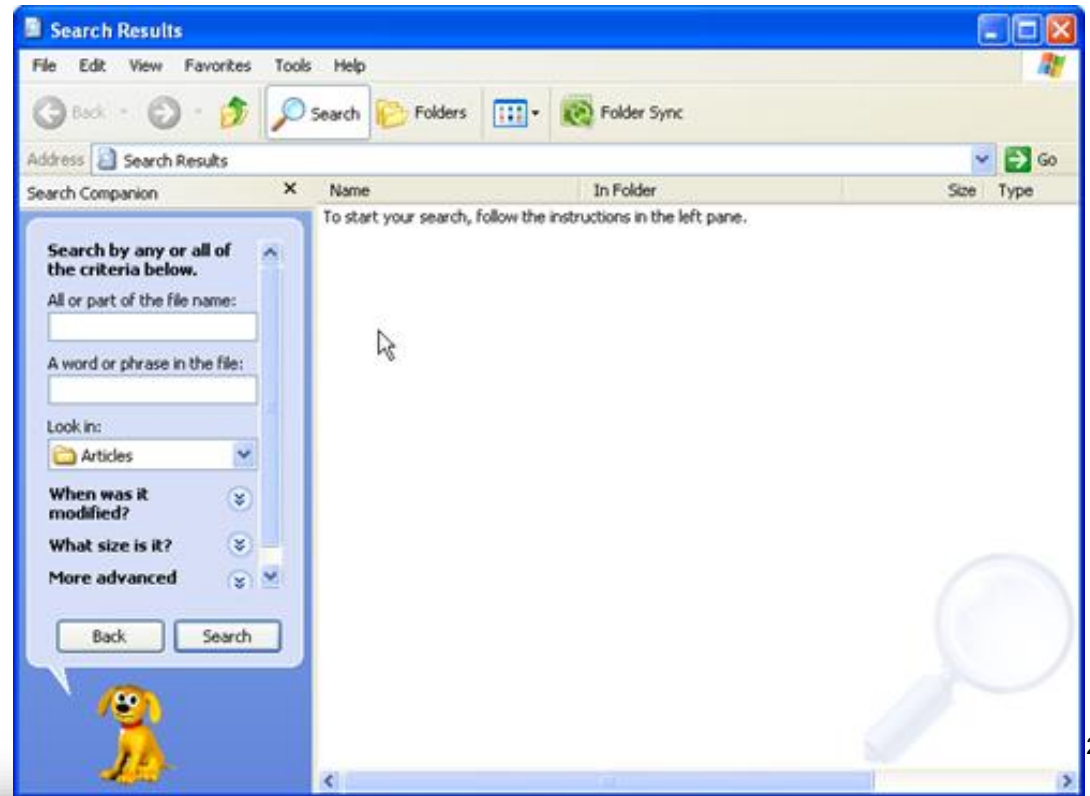
- ❑ Future directions for better integration:
 - ❑ NAS to provide fast diff from last index (Faster, prevents network RTT)
 - ❑ Move from “pull” to “push” (Improve SLA)



Do you always know path to your data?

NAS = Network Attached... Search?

- Search by content
- Search by file name
- Search by owner
- Search by type
- ...



Questions

