NameNode SPOF

Hadoop1.x

Primary NN

```
┌─────────────────────┐
│  ┌───────────────┐  │
│  │  fsimg @       │  │
│  └───────────────┘  │
│  ┌───────────────┐  │
│  │  edits 1-19   │  │
│  └───────────────┘  │
│  ┌───────────────┐  │
│  │  edit in prog │  │
│  └───────────────┘  │
│                     │
│  ┌───────────────┐  │
│  │  fsimg 19     │  │
│  └───────────────┘  │
└─────────────────────┘
```

Secondary NN

```
┌─────────────────────┐
│                     │
│                     │
│  ┌───────────────┐  │
│  │  edits ...    │  │
│  └───────────────┘  │
│  ┌───────────────┐  │
│  │  fsimg 19     │  │
│  └───────────────┘  │
└─────────────────────┘
```
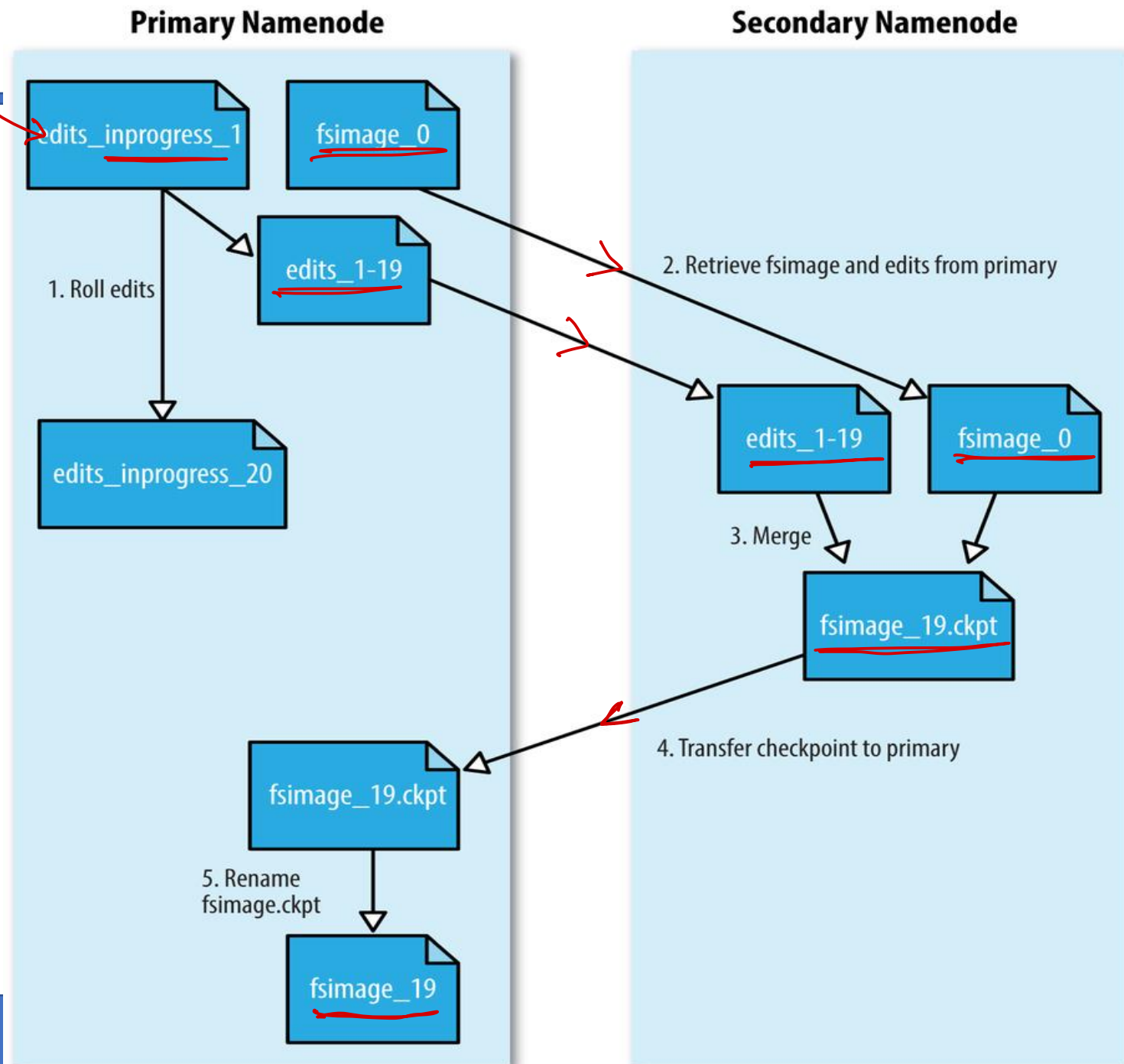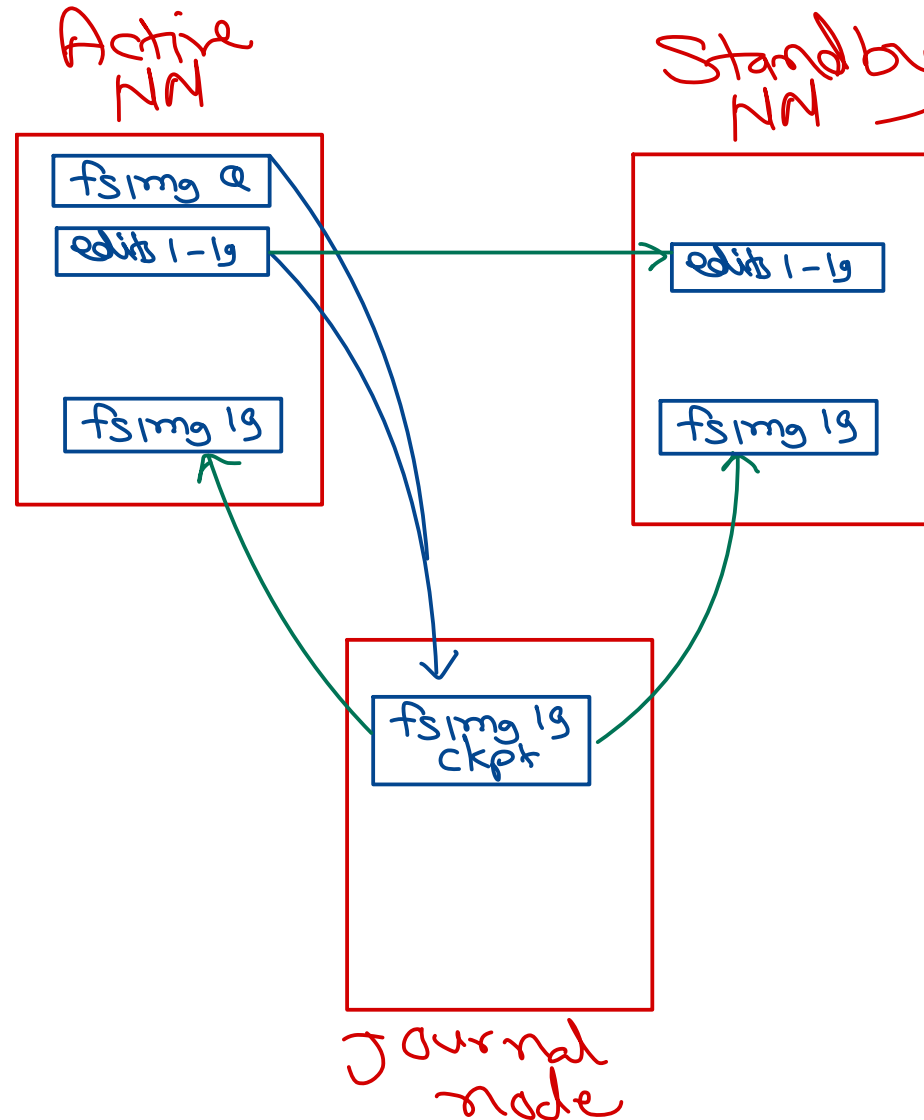
HDFS Internals

2.x

change in metadata

- NameNode loads all metadata into RAM from HDFS fsimage.
- Each change in metadata is fetched by SNN with last fsimage.
- Applying those changes in fsimage, SNN creates next fsimage checkpoint (.ckpt)
- This checkpoint is transferred to NameNode.
- NameNode rename it to consider as new fsimage and deletes old fsimage.
- This ensures that NN & SNN maintains same metadata.

Standby NameNode Hadoop2.x

HDFS - HA

Active NN

Standby NN

| fsimg Q |
| Edits 1-19 |

| fsimg 19 |

| Edits 1-19 |

| fsimg 19 |

| fsimg 19 ckpt |

Journal node

HDFS Java API

- HDFS can be accessed or manipulated using Java API.
- DistributedFileSystem class represent HDFS, while LocalFileSystem class represent local file system.
- Mainly two types of APIs
  - FileSystem API
  - File-IO API
- FileSystem API
  - Deals with metadata & directories.
  - FileStatus object contains metadata of file or directory.
  - Most of FileSystem APIs don't need access to DataNode (as metadata is maintained on NameNode itself).

- File IO API
  - Deals with data of the files.
  - FSDataInputStream class for reading the file, while FSDataOututStream class for writing the files.
  - They provide abstraction like replication process, network access, etc.
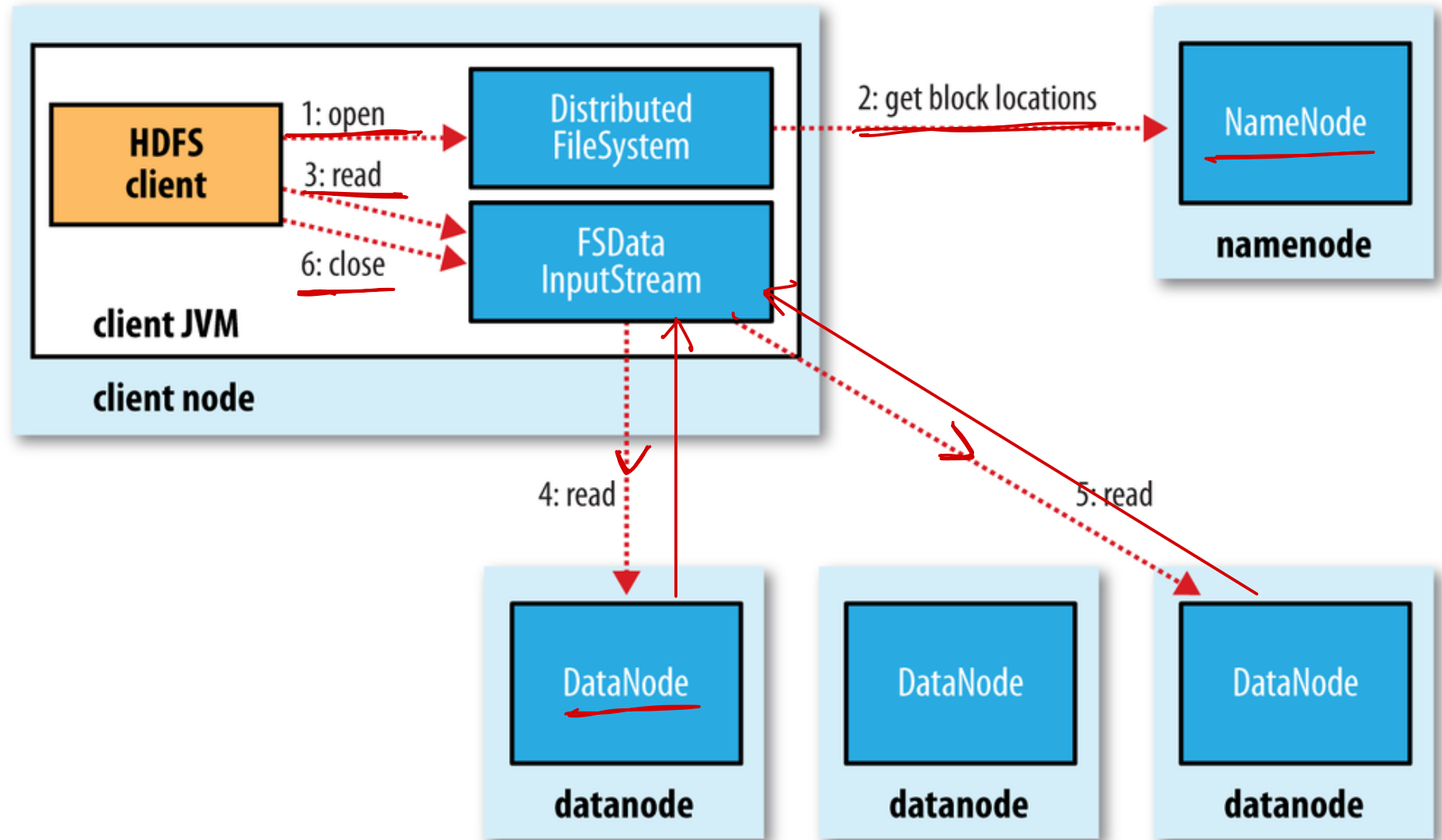
- Write/Read text files
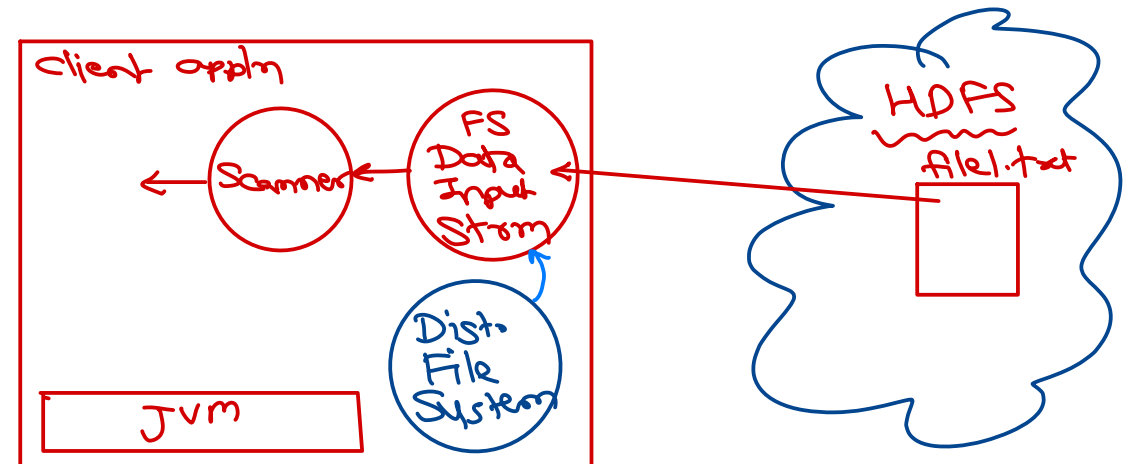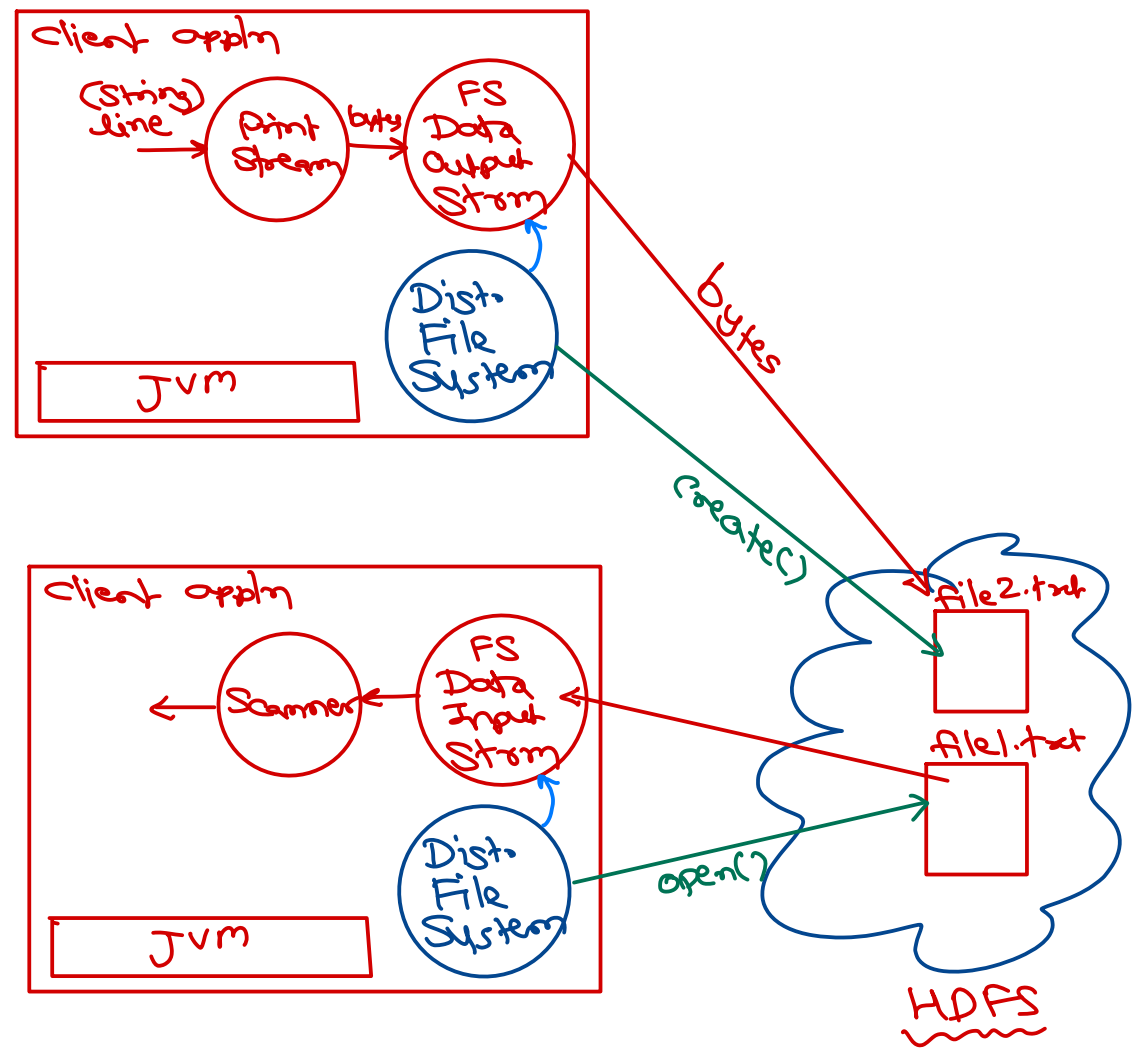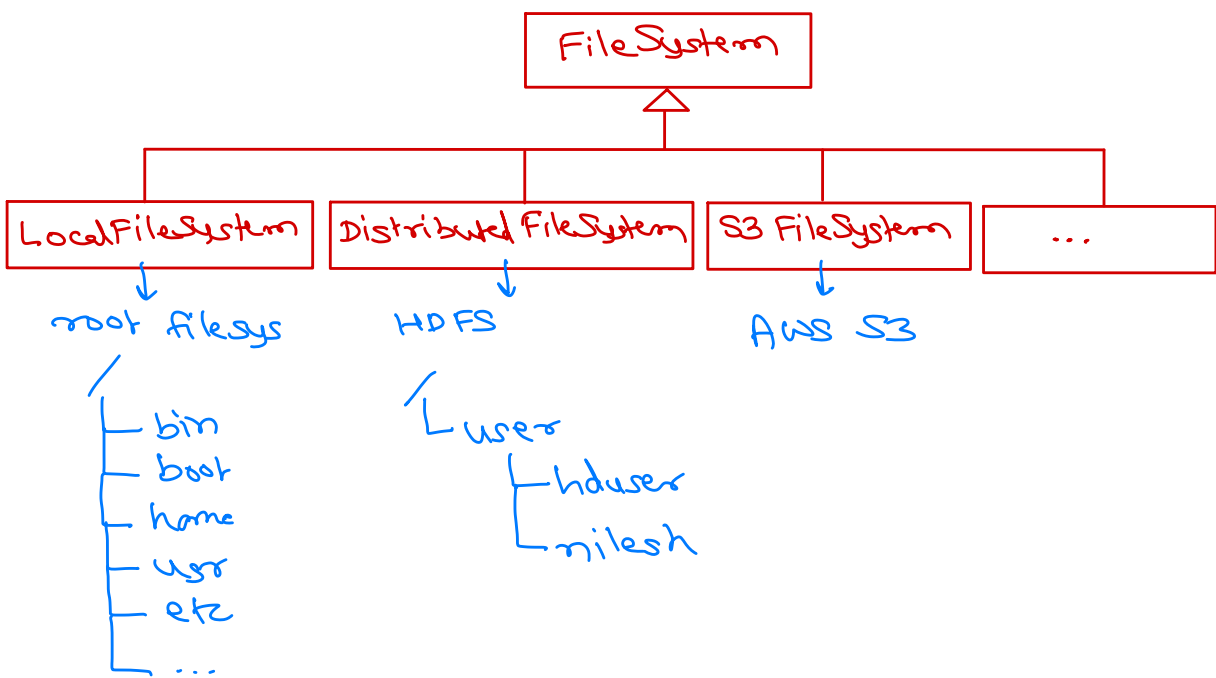
HDFS Java API: write operation
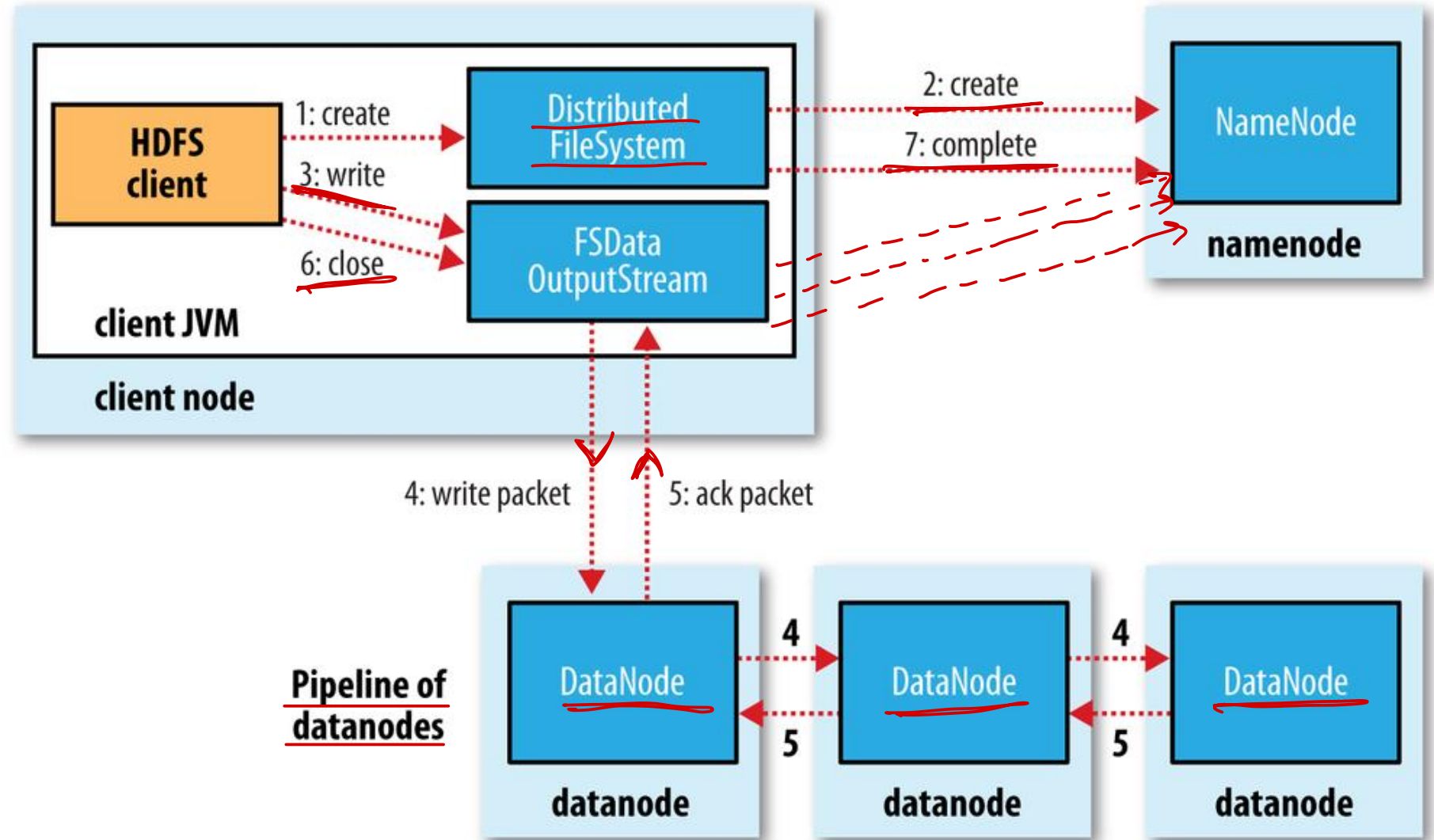
HDFS Java API: read operation

Thank you!

Nilesh Ghule <nilesh@sunbeaminfo.com>