Putting Things in Order
Effective Sorting

Share Orlando Spring 2025

Steve Pryor
Principal Technology Architect
precisely
steve.pryor@precisely.com

Sorting – An Essential Function

- Unsorted data is often not useful
- Order records in collating sequence
  - Usually EBCDIC, but maybe ALTSEQ (or EFS for DBCS)

- Sort is a *data transformation utility*
  - Filter and select input
  - Alter and create (multiple) outputs
  - Summarize and report

Sort is Everywhere

Sorting is Computer Science 101

- The Art of Computer Programming, Vol. 3, Knuth
- Internal sorting – bubble, shell, quicksort, heapsort, radix, many more
- External sorting – too much data to hold in memory

Commercial Sorts

- Precisely Syncsort MFX
- IBM DFSort
- Broadcom CA-Sort

Major Functions

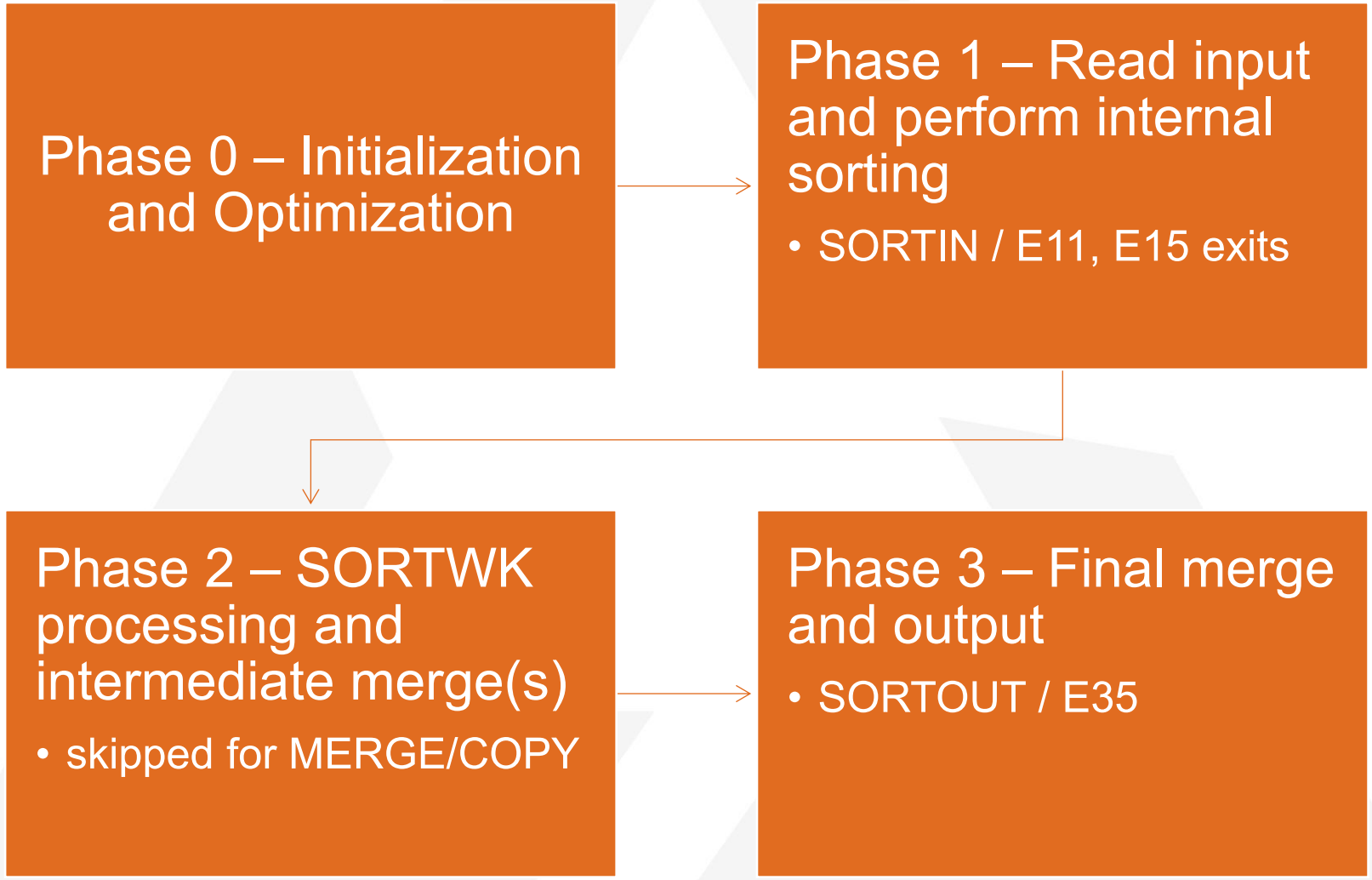
- SORT / MERGE / COPY
  - Put records in sequence order by *key* (or just copy),
    - SORT FIELDS=(p, l, f, s)
    - SORT FIELDS=(p, l, s),FORMAT=f
    - SORT FIELDS=COPY

```
SORT FIELDS=(5,4,PD,A,20,2,CH,D)
SORT FIELDS=(8,6,A),FORMAT=BI

SORT FIELDS=(7,3,CH,D,
             15.0,.1,D)
```

- Filter and modify records on input or output
  - INCLUDE / OMIT / INREC / OUTREC
- JOIN / REFORMAT
  - Merge two dissimilar files by key
- ICETOOL / SYNCTOOL
  - Reports and statistics

Sort Architecture and Flow of Control



Improving Performance - Limit the Input

- SKIPREC = nnnnn
  - Ignore the first nn records (from SORTIN / E15)

- INCLUDE / OMIT
  - Filter the input records on logical conditions
    - Fields within record
    - Field vs constant
    - Substring, bits on/off, numeric/alpha, date, et. al

```
OMIT COND=(7,3,ZD,EQ,18,2,PD,AND,12,2,BI,EQ,+25)
```

```
INCLUDE COND=(11,6,CH,EQ,C'HAMMER',OR,
              11,6,CH,EQ,C'CHISEL',OR,
              11,6,CH,EQ,C'SAW',OR,
              11,6,CH,EQ,C'WRENCH')
```

  - Only *one* INCLUDE / OMIT COND= statement is allowed for input
  - But *multiple* INCLUDE/OMIT= are allowed for OUTFIL (SORTOFxx) output

Improving Performance – Limit the Input

- E15 exit
  - Supply input or alter SORTIN input
  - Drop or shorten records

- INREC – reformat the input record
  - Modify, edit, and rebuild input record
    - Change columns, field formats, insert literals, much more
    - OVERLAY, PARSE, FINDREP, IFTHEN
  - Subsequent statements reference reformatted record
  - OUTREC will also reformat records, *after* Phase 2

```
INCLUDE COND=(5,1,GE,C'M'),FORMAT=CH
INREC FIELDS=(10,3,20,8,33,11,5,1)
```

Note: INREC can be used to convert ZD fields to PD, which sort faster

Improving Performance – Limit the Output

- Reduce the number / size of records
  - E35 to drop or alter output records
  - STOPAFT=nnnnn halt output after nn records written

- OUTREC – reformat record
  - Modify, edit, and rebuild output record
    - Reduce record size or *expand* after INREC processing

- Summarize duplicate-*sort-keyed* records
  - *Summary-key* fields added, duplicate record dropped
  - Specify EQUALS to keep first record
  - Specify NONE to keep *only* first record

```
SUM FIELDS=(5,5,ZD,12,
            6,21,3,
            35,7,ZD),FORMAT=PD
```

```
SUM FIELDS=NONE
```

Performance – Use Sort Resources Efficiently

- Give accurate filesizes
  - Specify SIZE= / FILESZ= if E15 supplies input (overestimate rather than under)

- Supply sufficient SORTWK (1.2 to 1.5x input size)
  - Allow sort to obtain via DYNALLOC parm
  - Increase DYNALLOC to 64 or provide JCL SORTWK for very large sorts
  - Never use tape SORTWK

- Provide sufficient virtual storage
  - Avoid small region sizes
  - Generally, allow the sort to default
    - Many options: MOSIZE, HSPRMAX, DSPSIZE, VSCORE, EMSPACE, MINCORE,et. al.

Use Multiple Output (OUTFIL)

Avoid multiple passes over the data

Many varying outputs from a single input

- OUTFIL FILES= multiple output files → SORTOFxx
  - ***Record Processing***
    - INCLUDE=/OMIT=, STARTREC/ENDREC
    - INREC=/OUTREC=
    - SPLIT/REPEAT many more
  - ***Report Writing***
    - HEADER/TRAILER
    - SECTIONS, counts, averages, totals, more
    - Generate XML or PDF output

```
OUTFIL FNAMES=OUT1,INCLUDE=(11,3,CH,EQ,C'D51'),ACCEPT=3
OUTFIL FNAMES=OUT2,STARTREC=2,ACCEPT=5

OUTFIL HEADER1=(2/,'Main Header Line',/,'Second Header',2/)
```

Use Human Resources Efficiently
Write Reports with SYNCTOOL / ICETOOL

- Batch utility invokes sort
  - Simplified control statements vs. OUTFIL
  - Creates multiple outputs/subsets
    - New records from multiple inputs
    - Count unique values, duplicates, headers, trailers
  - Formatted reports
    - Title, page no., date, headings, totals, etc

COPY – specify input / output
COUNT – no. of records in dataset
DATASORT – sort between header and trailer
DISPLAY – specify fields, headings, statistics
OCCUR – number of times each value occurs, numeric fields/dates
RANGE – range of values for numeric fields
SPLICE – join and match records

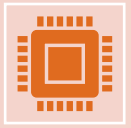
SMF TYPE-14 RECORDS 2021/08/04

DATE	TIME	SYS	JOBNAME	DATASETNAME
2025/08/04	18:48:45	S0W1	SJP1	SYS1.MACLIB
2025/08/04	18:48:47	S0W1	STEVE	STEVE.TEST.JCL
2025/08/04	18:48:55	S0W1	STEVE	STEVE.TEST.JCL
2025/08/04	18:49:13	S0W1	STEVE	S.SPFTEMP0.CNTL
2025/08/04	18:49:18	S0W1	STEVE	STEVE.TEST.JCL
2025/08/04	18:49:55	S0W1	STEVE	S.SPFTEMP0.CNTL
2025/08/04	18:49:55	S0W1	SJP2	SYS21216.T18495

Performance – Use Sort Resources Efficiently

Sort Accelerator instruction (SORTL)

Enabled by default



Global Dynamic Storage Manager (GDSM)

Monitors sort environment and adjusts virtual storage and SORTWK

- based on systemwide resource availability

ZPSaver – run sort in zIIP

Get help

Send sort SMF records for analysis

Summary

- Sort is more than sorting!
  - Copy, Merge, and Transform
  - Report Writing

- Performance matters!
  - Limit the number of input records (E15, SKIPREC, INCLUDE/OMIT)
  - Limit the amount of data to be sorted (INREC)
  - Limit the output (SUM, OUTREC)
  - Avoid multiple passes over the data (OUTFIL, SYNCTOOL/ICETOOL)
  - Use sort defaults and accelerators

QUESTIONS??

Your feedback is important!

SHARE
EDUCATE • NETWORK • INFLUENCE

Submit a session evaluation for each session you attend:

www.share.org/evaluation