Usenix FAST 2008 Conference, San Jose (02/27/2008)
Outrageous opinion statement

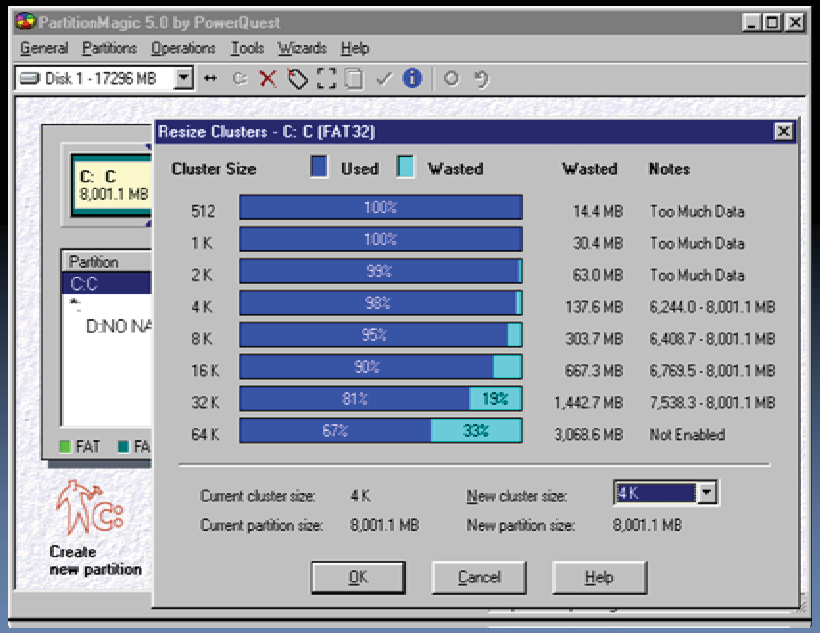
FILE SYSTEMS SHOULD BE LIKE BURGER MEALS: SUPERSIZE YOUR ALLOCATION UNITS !

Introduction

- Name: Konstantin Koll
- Age: 29
- Origin: University of Dortmund, Germany
- Occupation: PhD student
  - Topic: file systems
  - Created a high-performant relational file system (in a nutshell, a working clone of Microsoft WinFS)
  - During development, performance problems occured (one of them due to small allocation units)

Large allocation units are bad – are they ?

- Only full units can be allocated to a file, leading to wasted memory at the end
- Large allocation units ➔ large waste
- FAT file system got booted for clusters of 32 KB
- Tools exist to resize allocation units:

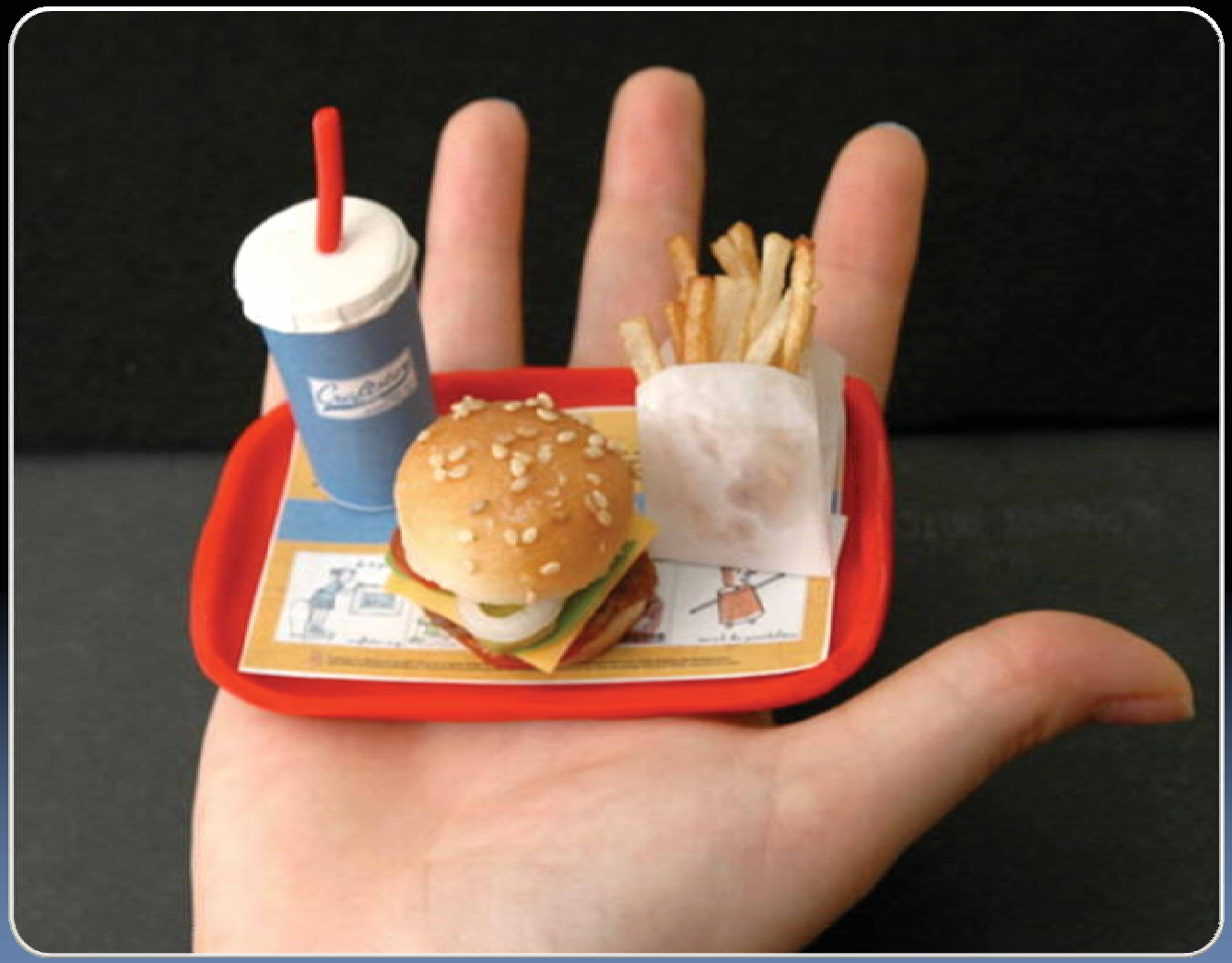


Burger allocation unit

- File system designers (you !) favor small allocation units to minimize wasted memory
- This is irresponsible !
- To make obvious why, let´s introduce the »BAU« (i.e. the size of a burger meal)

Small burger allocation units

- To reduce waste of food, let´s use small BAUs:



Small burger allocation units

- Benefit:
  - No waste, because only the required amount of food is being purchased by customers

Small burger allocation units

- Benefit:
  - No waste, because only the required amount of food is being purchased by customers
- Downsides:
  - Slow performance of food intake (customers have to go to the counter all the time)
  - High administrative overhead (during production, wrapping and delivery)
  - Small allocation units are pointless (food is cheap, so wasting some is irrelevant)

Small file allocation units

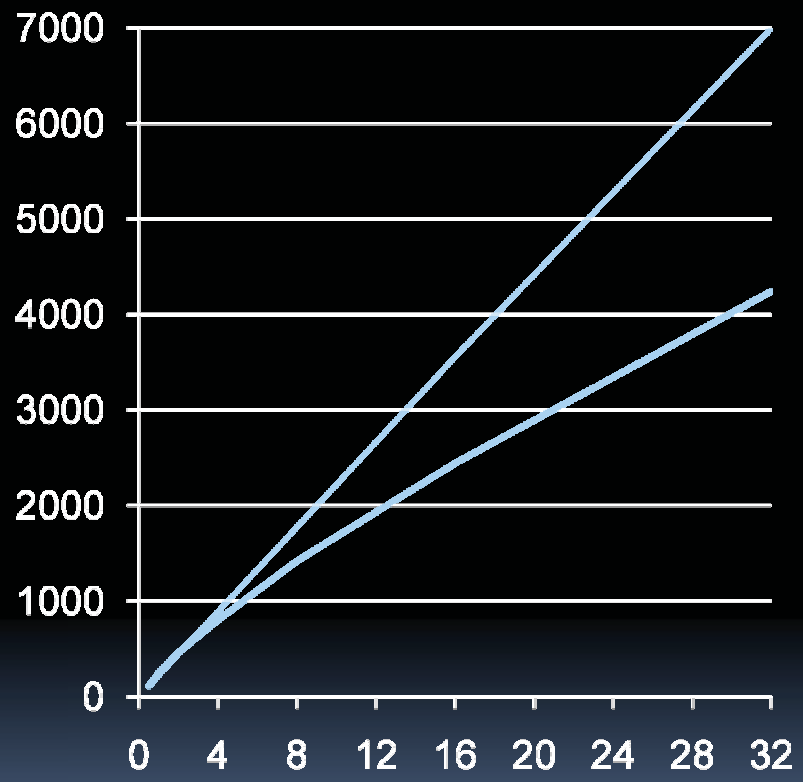
- Benefit:
  - No waste, because only the required amount of memory is being used by files

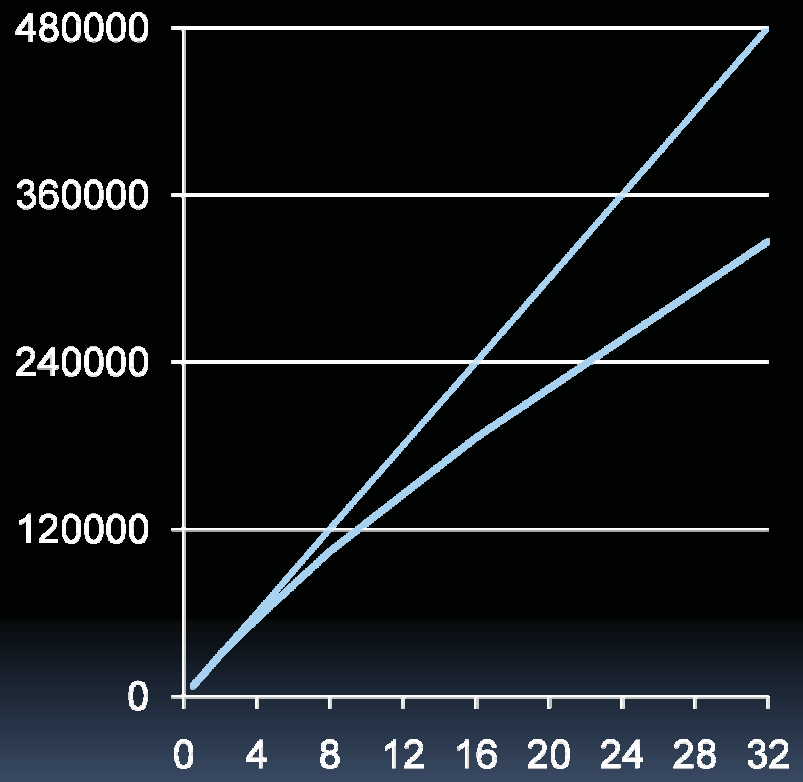
Small file allocation units

- Benefit:
  - No waste, because only the required amount of memory is being used by files
- Downsides:
  - Slow performance of I/O operations (cannot use burst reads for many continous sectors)
  - High administrative overhead (during file access)
  - Small allocation units are pointless (memory is cheap, so wasting some is irrelevant)

How much memory will really be wasted ?

- Study of actual user data (two examples):





  - Size of allocation units (X) vs wasted memory (Y)
  - Waste does **not** grow proportional to cluster size !

Conclusion

- Supersize your allocation units !