CU_CURR501
Page 1 of 5

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

July 25, 2024
9:41:00 AM

 Academic Org: Div of Computer Science & Engg – Subject: Computer Science

Course: CSCI5550　　　　**Course ID:** 011418　　　　**Eff Date:** 2024-07-01　　　　**Crse Status: Active**　　　　**Apprv. Status:** Approved　　　　**[**New Course**]**
Advanced File and Storage Systems 高級檔案及存儲系統

This course aims to introduce important systems-level research topics in the design and implementation of practical file and storage systems. Topics include: (i) storage device organization (e.g., disk drives, disk arrays, RAID, solid state drives), (ii) file system design (e.g., log-structured file systems, distributed file systems), (iii) data availability (e.g., erasure coding techniques, data integrity protection), (iv) data consistency (e.g., journaling techniques), (v) data compression (e.g., deduplication), (vi) benchmarking (e.g., I/O metrics, benchmarking tools), etc. Depending on the current research trends, the course also discusses the latest applied storage topics, especially related to scalable and dependable big data management.
Advisory: Students are expected to have taken CSCI3150 or ESTR3102 or equivalent.

本科旨在介紹有關設計和實踐實用檔案和存儲系統的重要系統研究課題。主題包括：（一）存儲設備結構（如磁盤驅動器、磁盤陣列、RAID、固態硬盤），（二）檔案系統的設計（如日誌結構檔案系統、分佈式檔案系統），（三）數據可用性（如糾刪編碼技術、數據完整性保護），（四）數據一致性（如日誌技術），（五）數據壓縮（如重複數據刪除），（六）基準評價（如I/ O指標、基準測試工具）等。本科也會按目前研究趨勢討論最新的應用存儲課題，尤其關於可擴展和可靠的大數據管理。
建議：學生應曾修讀CSCI3150或ESTR3102或同等學歷。

Grade Descriptor:　　　　A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

有關等級說明的資料，請參閱英文版本。

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

CU_CURR501
Page 2 of 5

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

July 25, 2024
9:41:00 AM

有關等級說明的資料，請參閱英文版本。

C

FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D

MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N

CU_CURR501
Page 3 of 5

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

July 25, 2024
9:41:00 AM

Course Attributes:
MSc Computer Science
MPhil-PhD Computer Sci & Erg

Topics:

<div align="center">

COURSE OUTCOMES

</div>

Learning Outcomes:

At the end of the course of studies, students will have acquired the ability to
1. Design, implement, evaluate, and deploy a practical storage system,
2. identify the research trends in file and storage systems, and
3. develop research skillsets in the area of file and storage systems.

Course Syllabus:

This course aims to introduce important systems-level research topics in the design and implementation of practical file and storage systems. Topics include: (i) storage device organization (e.g., disk drives, disk arrays, RAID, solid state drives), (ii) file system design (e.g., log-structured file systems, distributed file systems), (iii) data availability (e.g., erasure coding techniques, data integrity protection), (iv) data consistency (e.g., journaling techniques), (v) data compression (e.g., deduplication), (vi) benchmarking (e.g., I/O metrics, benchmarking tools), etc. Depending on the current research trends, the course also discusses the latest applied storage topics, especially related to scalable and dependable big data management.

Assessment Type:

Essay test or exam	: 30%
Others	: 40%
Presentation	: 30%

Feedback for Evaluation:

1. Quiz and examinations
2. Course evaluation and questionnaire
3. Question-and-answer sessions during class
4. Student consultation during office hours or online

Required Readings:

To be provided by course teacher.

Recommended Readings:

Reference academic papers:

1. C. Ruemmler and J. Wilkes. An Introduction to Disk Drive Modeling, IEEE Computer, 27(3):17-29, March 1994.

2. B. Schroeder and G. Gibson. Disk failures in the real world: what does an MTTF of 1,000,000 hours mean to you? Proceedings of USENIX Conference File and Storage Technologies, 2007.

3. M. Holland, G. A. Gibson, and D. P. Siewiorek. Architectures and Algorithms for On-Line Failure Recovery in Redundant Disk Arrays. ournal of Distributed and Parallel Databases, Vol. 2, No. 3, 1994.

4. Chen, P. M., Lee, E. K, Gibson, G. A., Katz, R. H., Patterson, D. A. RAID: High-Performance, Reliable Secondary Storage. ACM Computing Surveys, 26(2):145-185, 1994.

5. N. Agrawal, V. Prabhakaran, T. Wobber, J. D. Davis, M. Manasse, and R. Panigrahy. Design Tradeoffs for SSD Performance. In Proceedings of USENIX Annual Technical Conference (ATC), 2008.

6. M. Rosenblum and J. K. Ousterhout. The Design and Implementation of a Log-structured File System. ACM Transactions on Computer Systems, 10:26–52, 1992.

7. S. Ghemawat, H. Gobioff, and S.-T. Leung The Google File System. Proceedings of ACM Symposium on Operating Systems Principles (SOSP), 2003.

8. J. S. Plank. A tutorial on Reed-Solomon coding for fault-tolerance in RAID-like systems. Software, Practice & Experience, 27 (9):995–1012, 1997.

9. G, Sivathanu, C. P. Wright, and E. Zadok. Ensuring data integrity in storage: techniques and applications. Proceedings of the 2005 ACM workshop on Storage security and survivability, 2005.

10. V. Prabhakaran, A. C. Arpaci-Dusseau, and R. H. Arpaci-Dusseau. Analysis and evolution of journaling file systems. Proceedings of the USENIX Annual Technical Conference (ATC), 2005.

11. B. Zhu, K. Li, and H. Patterson. Avoiding the disk bottleneck in the data domain deduplication file system. Proceedings of USENIX Conference File and Storage Technologies, 2008.

Reference book:

12. D. Giampaolo. Practical File System Design with the Be File System. Morgan Kaufmann Publishers, 1999.

OFFERINGS	
1. CSCI5550	Acad Organization=CSEGV; Acad Career=RPG

COMPONENTS
LEC : Size=30; Final Exam=Y; Contact=3
TUT : Size=30; Final Exam=N; Contact=1

ENROLMENT REQUIREMENTS	
1. CSCI5550	**Enrollment Requirement Group:**
	For undergraduate students in Computer Science (CSCIU and CSCIN) or Computer Engineering (CENGU and CENGN); or
	For students in MSc Computer Science or MPhil-PhD Computer Science & Engineering

CU_CURR501
Page 5 of 5

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

July 25, 2024
9:41:00 AM

Additional Information

VTL-Onsite face-to-face hrs   0
VTL-Online synch. hrs   0
VTL-Online asynch. hrs   0

< E N D   O F   R E P O R T >