

Archiving Dynamic Websites

at ibiblio.org

Me

Cristóbal Palmer

cmp@cmpalmer.org

@coxn



81A2 0F2F 9C11 79B2 03AB 0E1B 48C0 5B6D DD36 6EEB

You

- Shirt?
- Run a website of any kind whatsoever?
- Looked for something on the web and found it broken or missing?
- Library card?

Overview

- Archives
- Digital Archives
- ibiblio
- Archiving the Web
- Archiving Dynamic Sites at ibiblio
- Join us

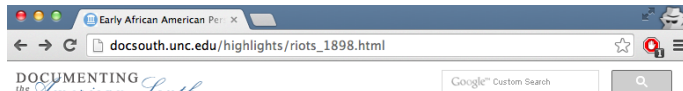
Archives



Photo by William Yeung <http://flic.kr/p/foaQ3L> CC-BY 2.0

Digital Archives

Digital Archives



DOCUMENTING
the
American South

[About](#) | [Collections](#) | [Authors](#) | [Titles](#) | [Subjects](#) | [Geographic](#) | [K-12](#) | [Facebook](#) | [Buy DocSouth Books](#)

Collections >> Highlights >> Early African American Perspectives on the Wilmington Race Riots of 1898

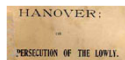
HIGHLIGHTS

Early African American Perspectives on the Wilmington Race Riots of 1898

On Thursday, November 10, 1898, Colonel Alfred Moore Waddell, a Democratic leader in Wilmington, North Carolina mustered a white mob to retaliate for a controversial editorial written by Alexander Manly, editor of the city's black newspaper, the *Daily Record*. The mob burned the newspaper's office and incited a bloody race riot in the city. By the end of the week, at least fourteen black citizens were dead, and much of the city's black leadership had been banished. This massacre further fueled an ongoing statewide disfranchisement campaign designed to crush black political power. Contemporary white chronicles of the event, such as those printed in the *Raleigh News and Observer* and Wilmington's *The Morning Star*, either blamed the African American community for the violence or justified white actions as necessary to keep the peace. African American writers produced their own accounts—including fictional examinations—that countered these white supremacist claims and highlighted the heroic struggles of the black community against racist injustice.

J. Allen Kirk's "A Statement of Facts Concerning the Bloody Riot in Wilmington, N.C. Of Interest to Every Citizen of the United States," seeks to "impartially" recount the event for both white and black audiences. Kirk, the pastor of a Wilmington church, refers to himself almost exclusively in the third person and explains that he writes about the massacre in hopes that he might "prevent exaggerated stories from being heaped upon the community or upon the country." He describes hiding with his family in the African American cemetery outside of town, hearing about the burning of the *Record* office, and watching as a mob of white men escorted white women and children to safety while black citizens fled to the swamps or woods to avoid being killed in the streets. Even as late as the following Sunday, the mob was still searching for missing black leaders, including Kirk, who was eventually forced to leave his

Other Highlights



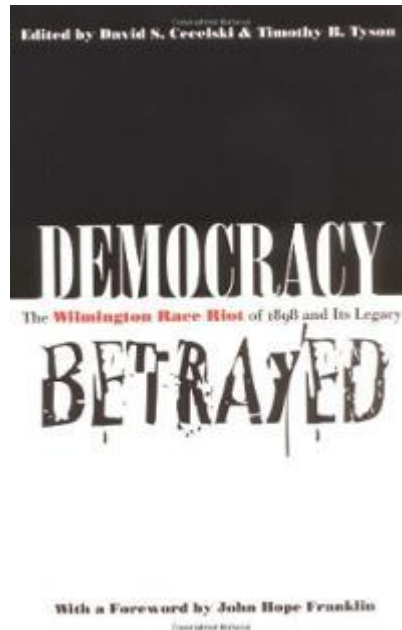
Title Page, Fulton's *Hanover*

Also of Interest:

Durham business leader John Merrick's speech about the riots (p. 161)

A guide to Wilmington published 4 years after the riots

A collection of articles written by David Bryant Fulton: "Eagle Clippings"



digital archiving challenges

- compared to archival paper, digital media does not last

http://static.googleusercontent.com/media/research.google.com/en/us/archive/disk_failures.pdf

Failure Trends in a Large Disk Drive Population

Eduardo Pinheiro, Wolf-Dietrich Weber and Luiz André Barroso

Google Inc.

1600 Amphitheatre Pkwy

Mountain View, CA 94043

{edpin,wolf,luiz}@google.com

Abstract

It is estimated that over 90% of all new information produced in the world is being stored on magnetic media, most of it on hard disk drives. Despite their importance, there is relatively little published work on the failure patterns of disk drives, and the key factors that affect their lifetime. Most available data are either based on extrapolation from accelerated aging experiments or from relatively modest sized field studies. Moreover, larger population studies rarely have the infrastructure in place to collect health signals from components in operation, which is critical information for detailed failure analysis.

We present data collected from detailed observations of a

for guiding the design of storage systems as well as devising deployment and maintenance strategies.

Despite the importance of the subject, there are very few published studies on failure characteristics of disk drives. Most of the available information comes from the disk manufacturers themselves [2]. Their data are typically based on extrapolation from accelerated life test data of small populations or from returned unit databases. Accelerated life tests, although useful in providing insight into how some environmental factors can affect disk drive lifetime, have been known to be poor predictors of actual failure rates as seen by customers

digital archiving challenges

- copy-on-write, checksums, etc. (ZFS, btrfs)
- LOCKSS, iRODS, etc.

ibiblio

ibiblio.org

The public's library
and digital archive

1. sils.unc.edu
2. jomc.unc.edu
3. its.unc.edu



ibiblio past

sunsite.unc.edu

metalab.unc.edu

ibiblio.org



IT'S BELIEVED
THAT IN 1994 WXYC
IN CHAPEL HILL
IN THIS STATE
BECAME THE FIRST
RADIO STATION
TO WEBCAST

ibiblio past

- HTML
- dynamic in a different sense



Welcome!

Click on a Topic!



**Halloween
Traditions**



**Costume
Shopping**

**Click
to See
Halloween
Night
in
Chapel Hill,
NC**



**Scary
Stories**



**Haunted
Houses**

#SMTIMFH

#\$-:Image Map file created by Map THIS!

#\$-:Map THIS! free image map editor by Todd C. Wilson

#\$-:Please do not edit lines starting with "\$"

#\$VERSION:1.30

#\$DATE:Tue Oct 29 20:36:32 1996

#\$PATH:C:\temp\

#\$GIF:secondpage4.gif

#\$FORMAT:nca

#\$EOH

Format: NCSA

rect Halicon.htm 14,309 189,433

rect tradit.htm 1,1 134,144

rect costume.htm 297,0 434,152

rect chapel.htm 159,132 297,257

rect witch.htm 3,232 125,383

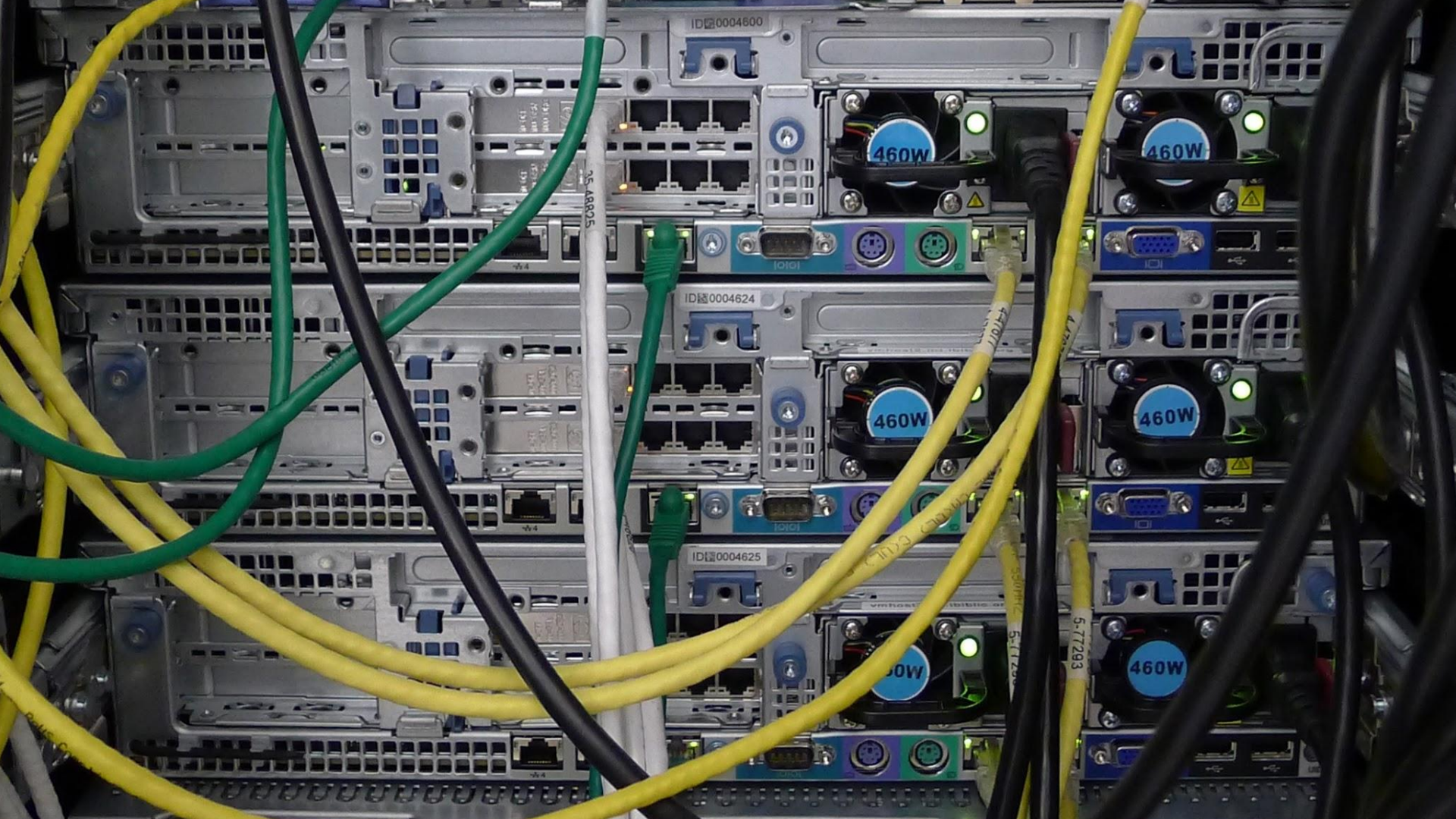
rect haunt.htm 318,244 416,384

default Halicon.htm

ibiblio present

- vmware
- even more dynamic
- expensive to support

Pause for some sites



this talk is a redux

When Good Sites Go Bad - And how to save them: Archival challenges and solutions for post-1995 websites

Paul Jones (director of ibiblio): <http://bit.ly/Qs4v0d>

Archiving the Web

Archiving the Web

Wayback Machine is awesome!



Photo by Kent Kanouse <http://flic.kr/p/5KEqjP> CC-BY 2.0

Archiving Dynamic Sites

ibiblio future

- Free and Open Source hypervisor
- Free and Open Source Virtual Machine management layer
- metadata framework for collections and snapshots
- “checkout” mechanism

ibiblio future - hypervisor

Xen or KVM? Other?

ibiblio future - node mgmt

- Managing nodes
- Xen4CentOS
- oVirt
- libvirt

ibib future - magical cloud

- Manage apps/collections, not nodes
- OpenStack?? <http://www.openstack.org/software/>
- OpenShift?? <http://openshift.github.io/> (+ Docker <http://www.docker.io/>)

?

This talk:

<http://bit.ly/X4ft1L>

is licensed under
CC-BY-SA 3.0 US

http://creativecommons.org/licenses/by-sa/3.0/us/deed.en_US