Characterization and Mitigation of Failures in Complex Systems

System Dependability

- Dependability
  - Impairments
    - Faults
    - Errors
    - Failures
  - Means
    - Procurement
      - Fault prevention (FP)
      - Fault tolerance (FT)
    - Validation
      - Fault removal (FR)
      - Fault forecasting (FF)
  - Attributes
    - Reliability
    - Availability
    - Security
    - Safety

Reliability, Availability + Performance → Performability

Reliability, Availability + Service → Survivability

Security → Survivability

System Dependability

FP { Minimal maintenance

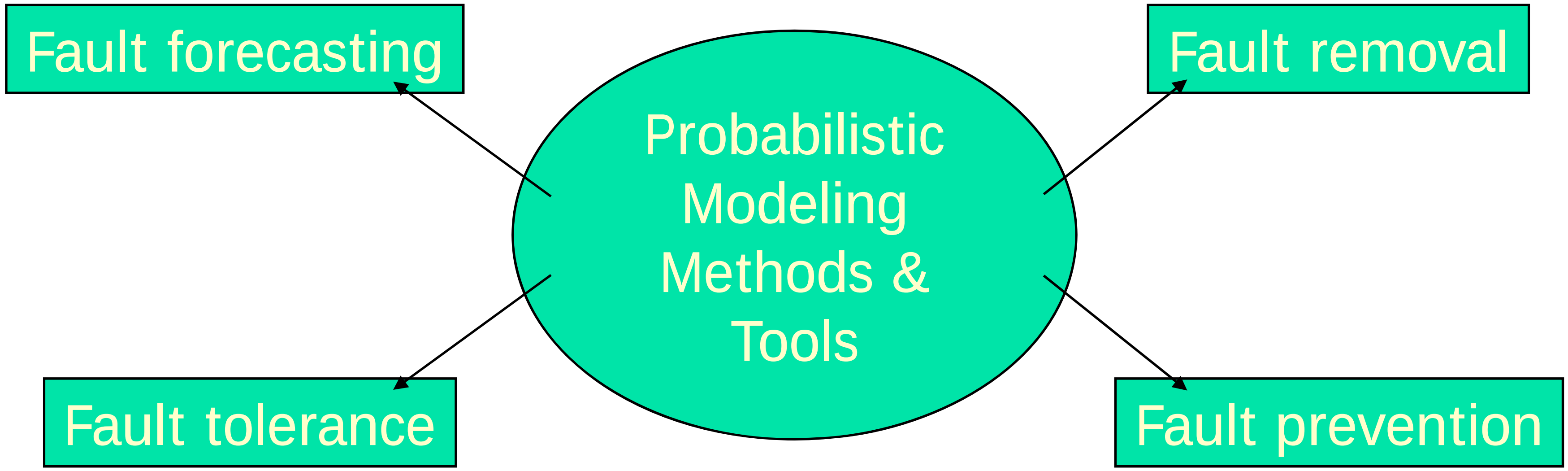
FT { Redundancy: hardware,software,information,time
Diversity: data, design, environment

FR { Verification (testing)
Maintenance: corrective (repair) and preventive

FF { Probabilistic modeling
Non-probabilistic modeling

} Applications

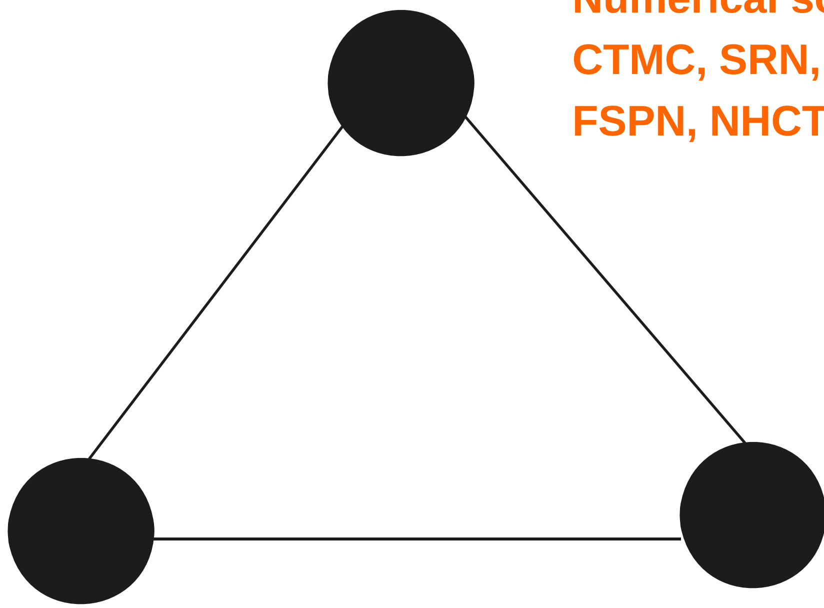
Modeling and Analysis



A Research Triangle

Theory

Modeling methods and
Numerical solutions:
CTMC, SRN, MRSPN
FSPN, NHCTMC, FTREE



Tools

SHARPE
SPNP
SREPT
SRA

Applications

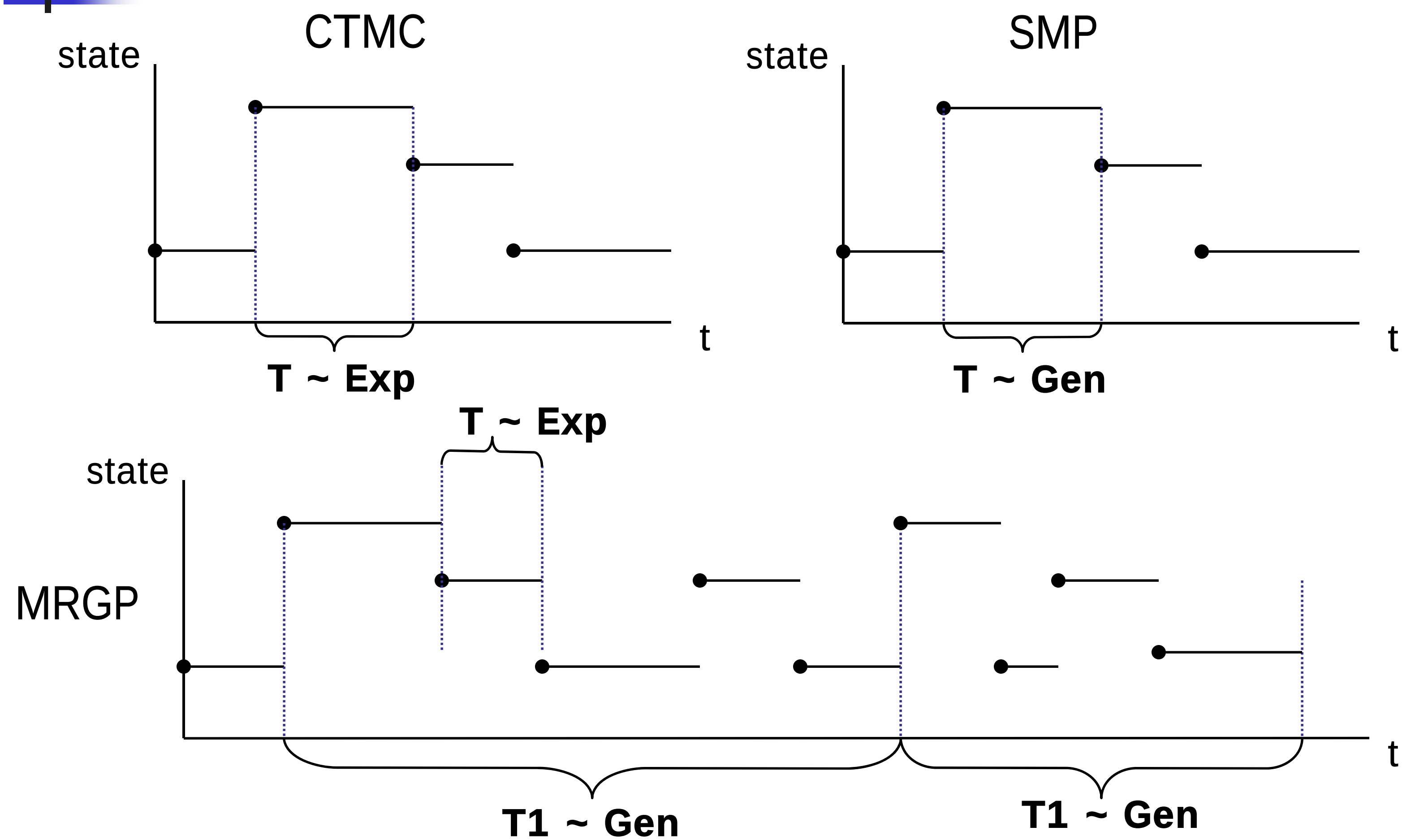
Hardware and software:
Fault-tolerant Systems
Real-time Systems
Computer Networks
Wireless Networks

Modeling Methods

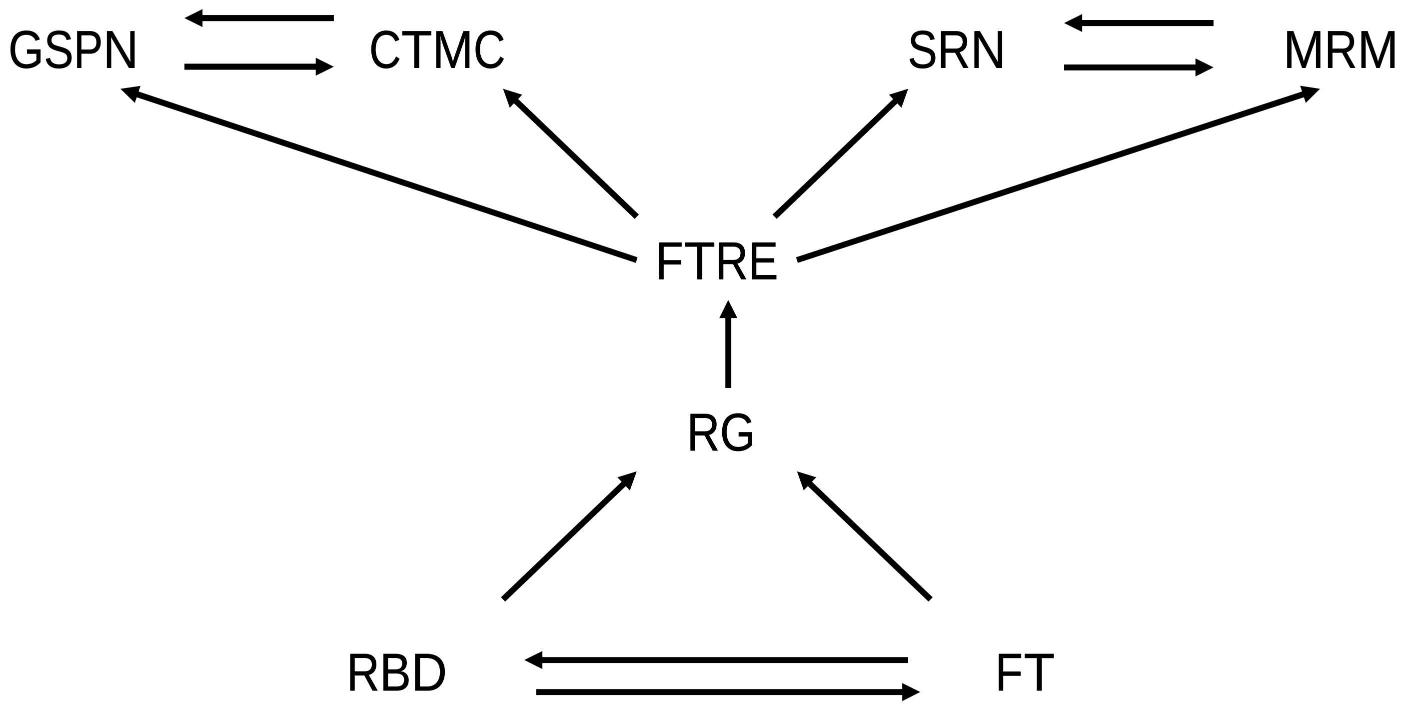
Probabilistic modeling
- Fault trees
- DTMC, CTMC → SMP → MRGP
- DTMC, CTMC → NHCTMC
- SPN → SRN
- SPN → MRSPN
- SPN → FSPN
- SDE
- FBM, FGN
- HS

Non-probabilistic modeling
- Nonlinear dynamics
- HS

CTMC → SMP → MRGP

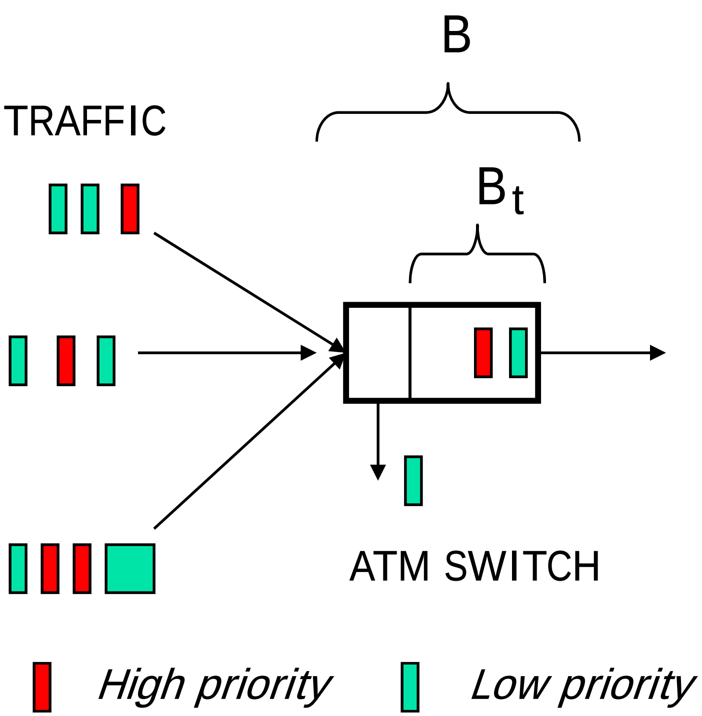


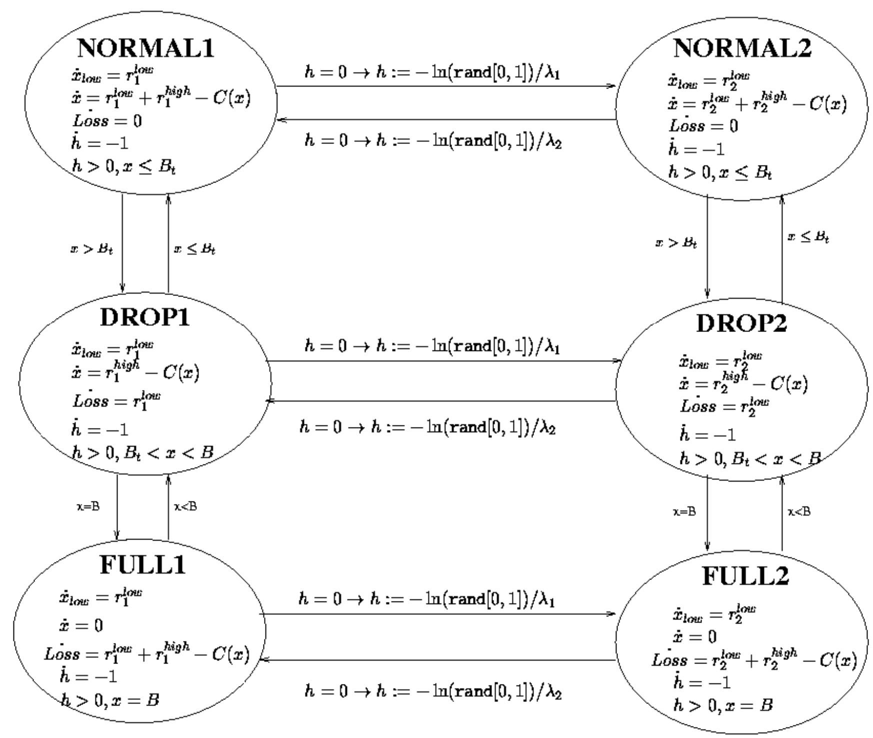
Relation: Dependability Models



Modeling of HS and FSPN

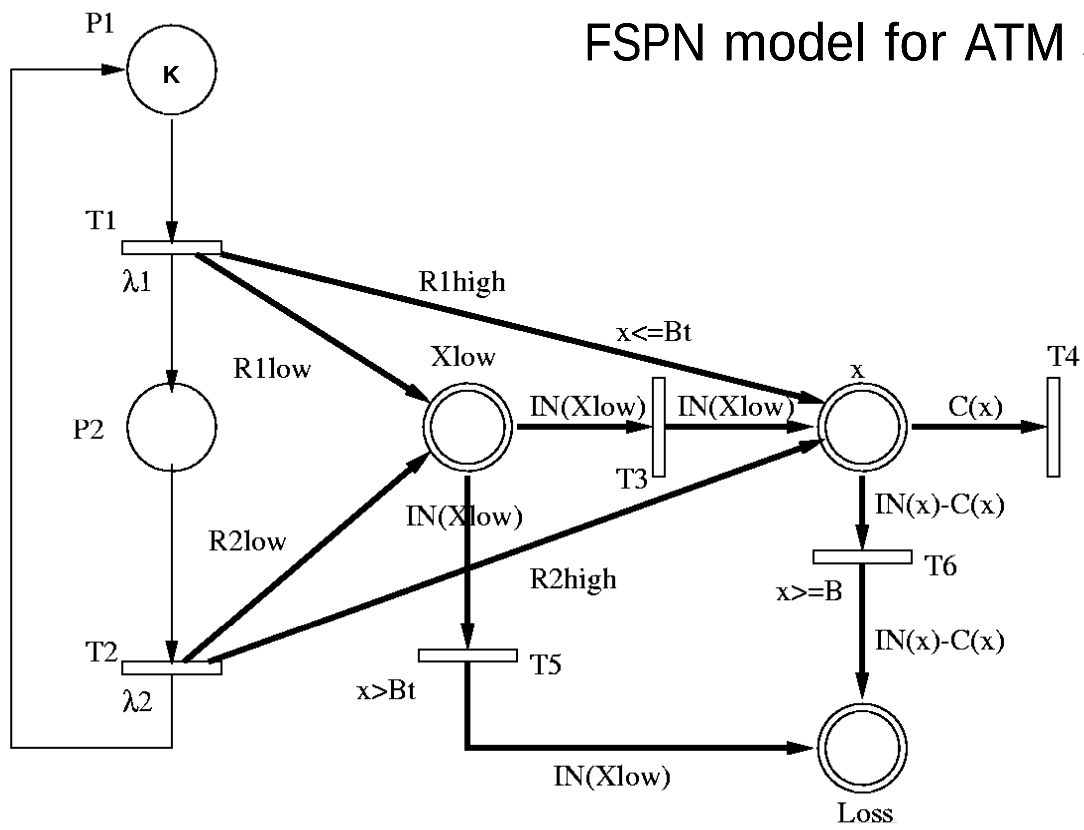
- Statistical multiplexed ATM switch and HS model:





Modeling of HS and FSPN

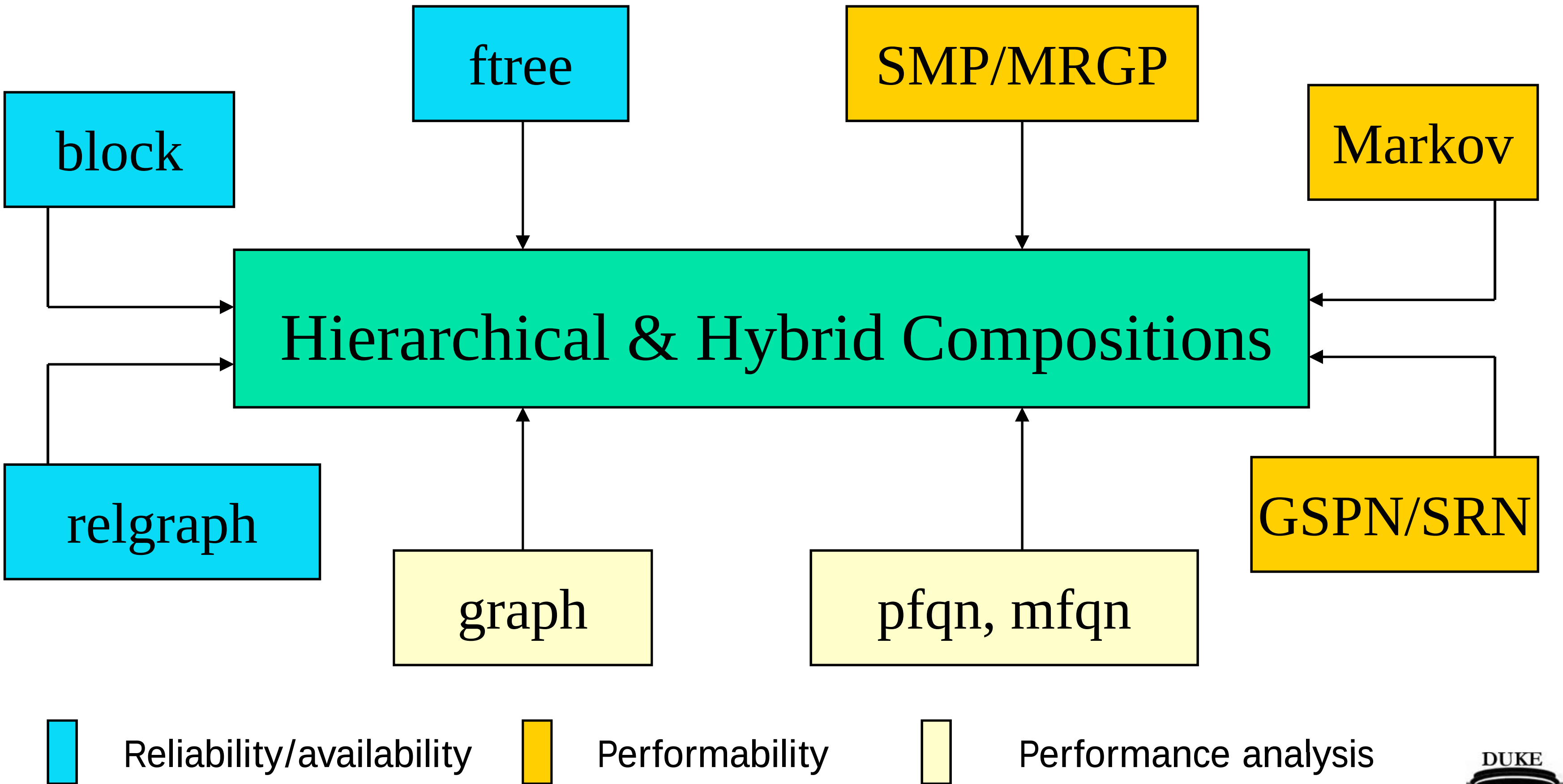
FSPN model for ATM switch



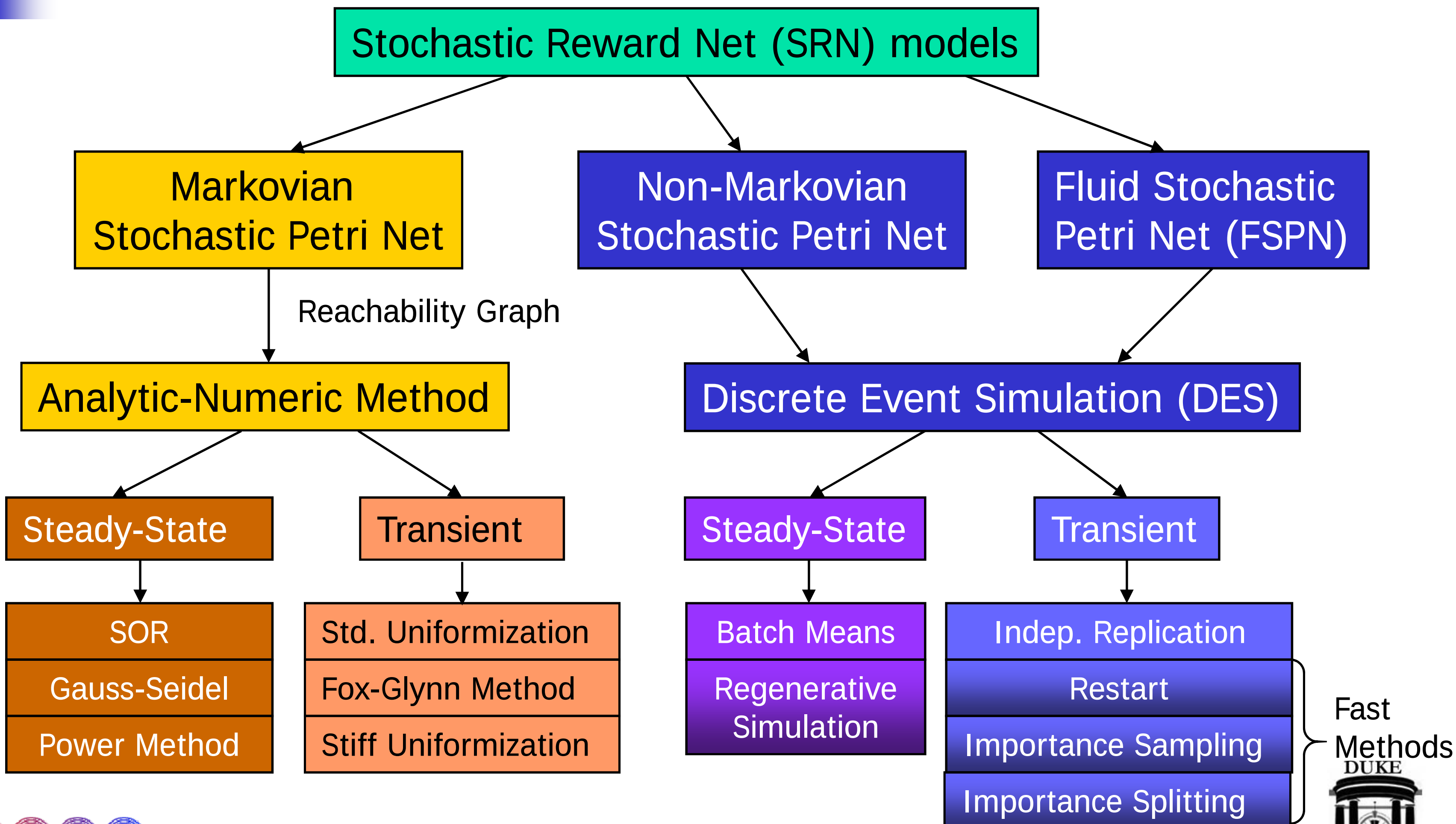
Solution Techniques

- Algorithms
  - Fast algorithms for network reliability and fault tree computation
  - Fast algorithms for SRN
- Numerical and simulation tools
  - SHARPE (symbolic hierarchical automated reliability and performance evaluator)
  - SREPT (software reliability estimation prediction tool)
  - SPNP (stochastic Petri net package)
  - SRA (software rejuvenation agents)

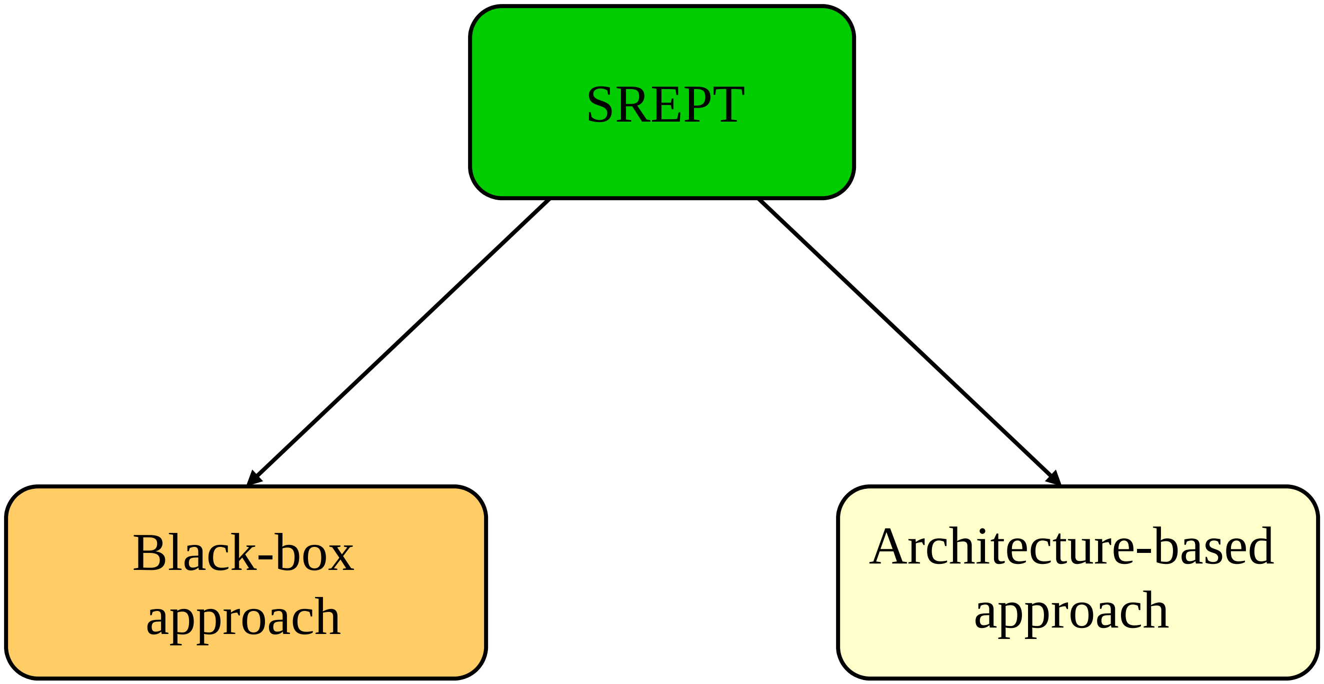
Architecture of SHARPE



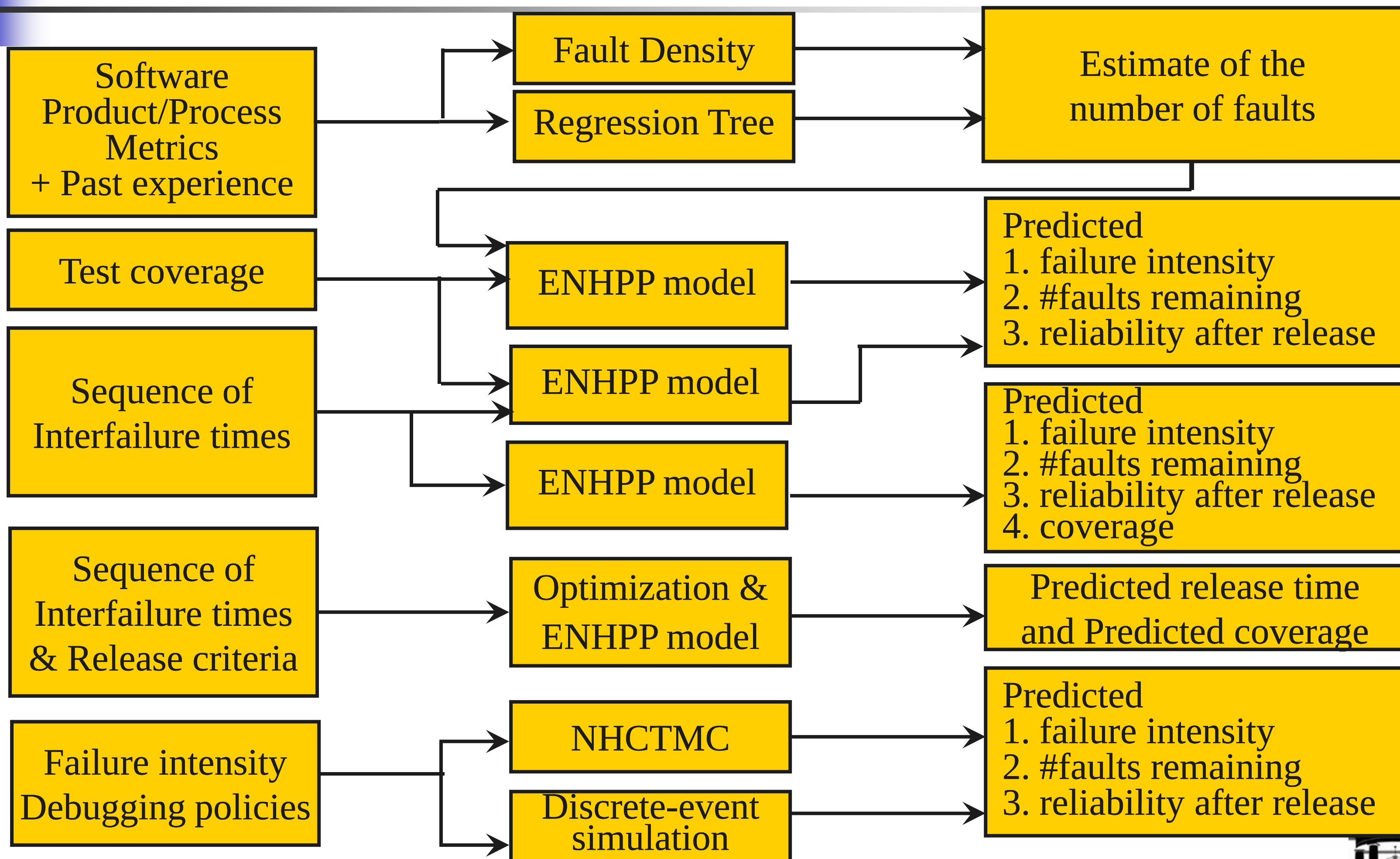
Architecture of SPNP



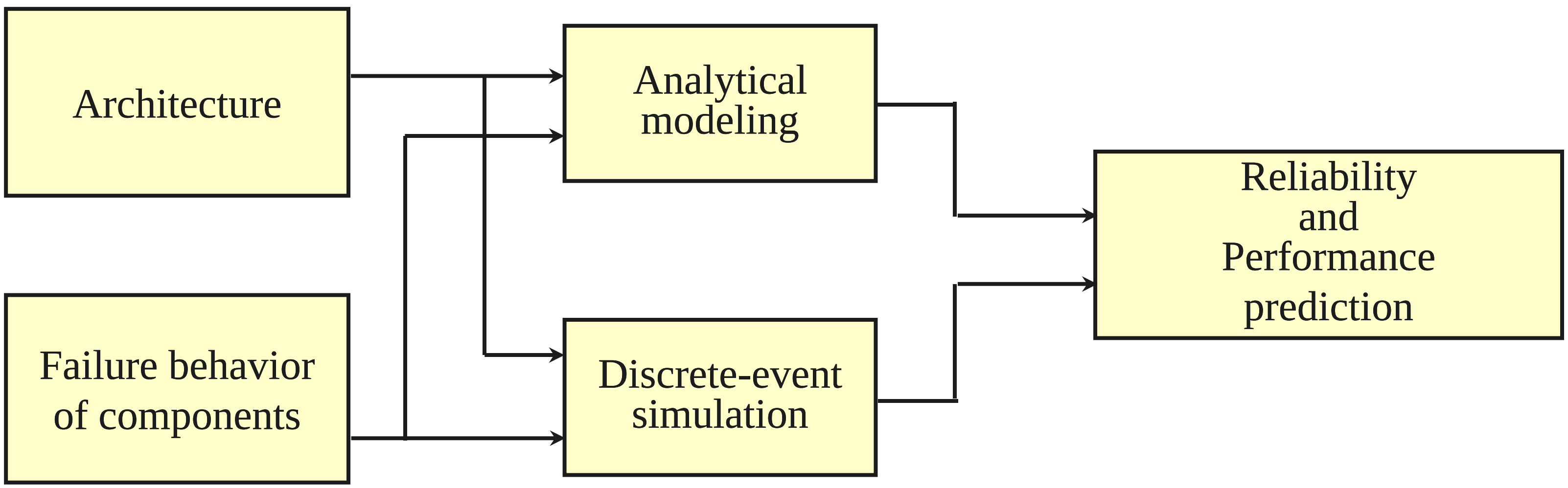
Architecture of SREPT



Black-box Approach



Architecture-based Approach



Challenges

- Largeness
  - Fault trees using BDD
  - Markov model
    - Automated generation/solution based SPN, SRN
    - Hierarchical methods, fixed point iteration
- Stiffness
- Non-exponential distributions
  - NHCTMC
  - SMP, MRGP
  - DES (Discrete-Event Simulation)
- Combined performance and reliability/availability —— MRM ——→ SRN
- Joint continuous & discrete state spaces
  - FSPN
  - HS
  - SDE

Applications

- For high-performance and high-dependability
- Software
  - Software reliability (FR,FF)
  - Software rejuvenation (FF,FR)
  - Software fault-tolerance (FT,FF)
- Hardware (with software)
  - Preventive maintenance (FR,FF)
  - Availability model (FF)
  - Upgrade design (FT,FF)
  - Performability of wireless communications (FF)
  - Transient behavior of ATM networks with failure (FF)
  - Error recovery in communication networks (FF)

Software Reliability

- Early prediction of quality based on software product/process metrics.
- Reliability growth modeling based on failure and coverage data collected during testing
- Architecture-based software reliability
- Developed a tool (SREPT)

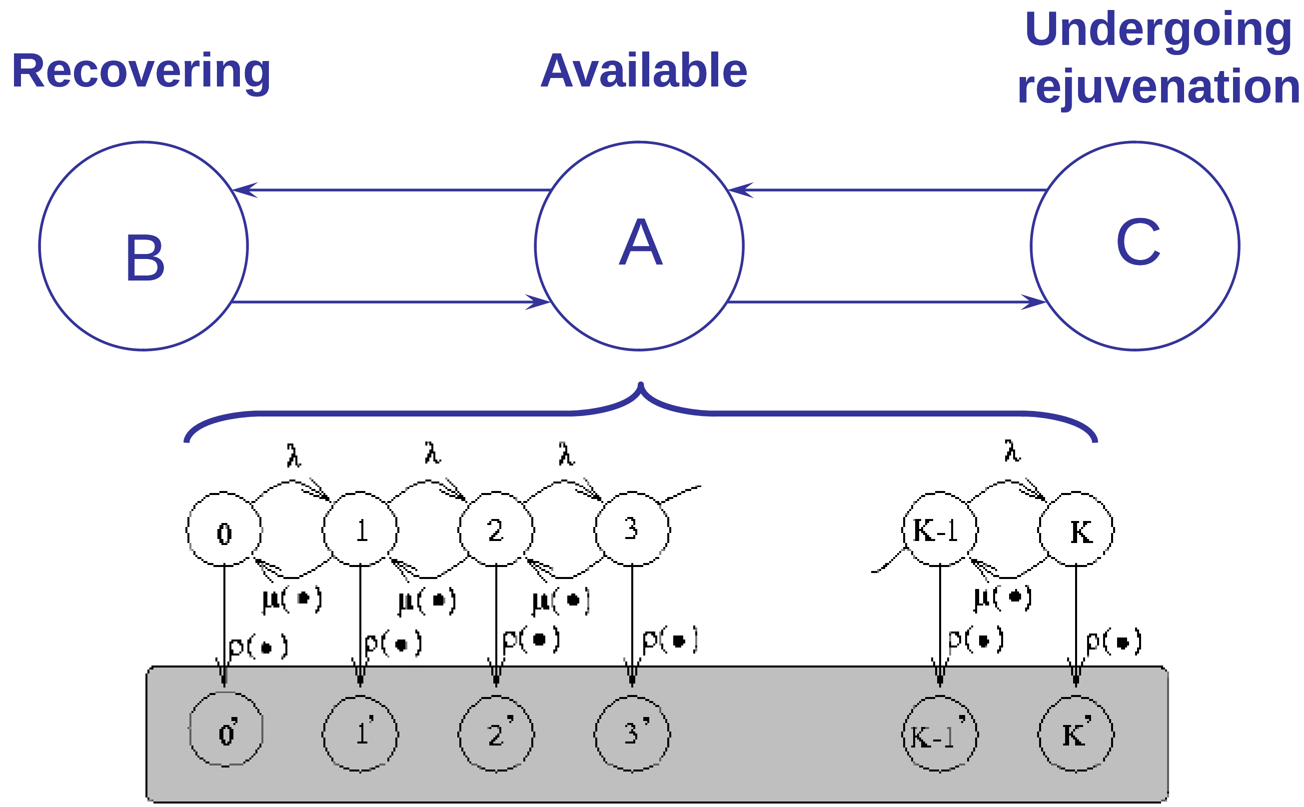
Software Rejuvenation

Definition: proactive fault management technique for the systems to counteract the effect of aging

Approaches:

- Periodic: Rejuvenation at regular (deterministic) time intervals irrespective of system load
- Periodic and instantaneous load: Rejuvenation at regular intervals and wait until system load is zero
- Prediction-based: Rejuvenation at times based on estimated times to failure (implemented in IBM Netfinity Director)
  - Purely time-based estimation
  - Time and workload-based estimation

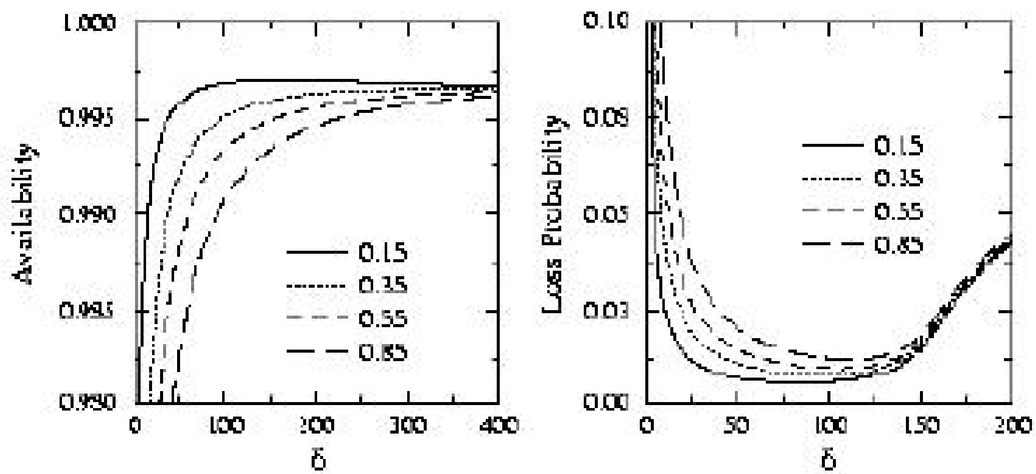
Analytical Modeling



Subordinated non-homogeneous CTMC for $t = \delta$

Numerical Example

Service rate and failure rate are functions of time, $\mu(t)$ and $\rho(t)$



Measurement-based Estimation

- Objective

  detection and validation of software aging

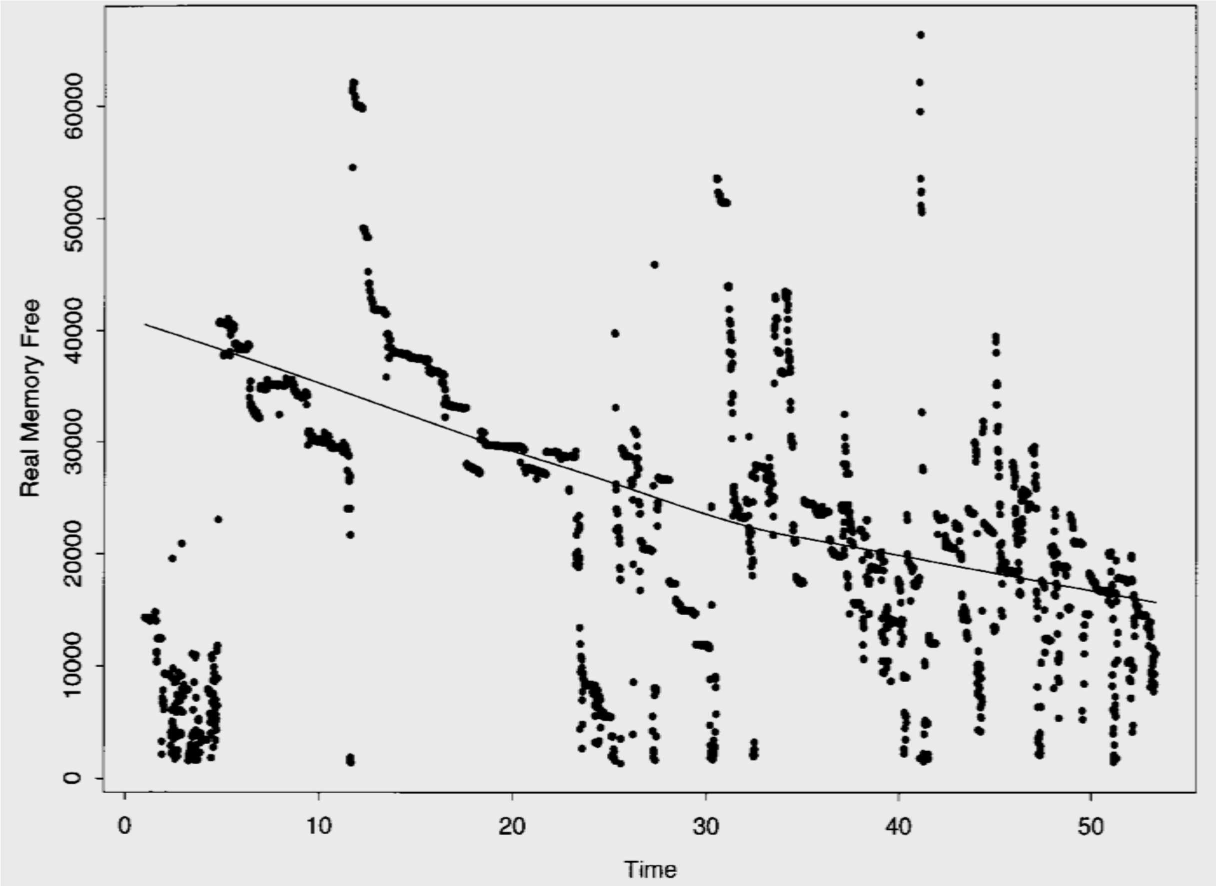
- Basic idea

  periodically monitor and collect data on the attributes responsible for determining the health of the executing software

- Quantifying the effect of aging

  proposed metric - *Estimated time to exhaustion*

Non-**parametric** Regression Smoothing of Rossby data - Real Memory Free

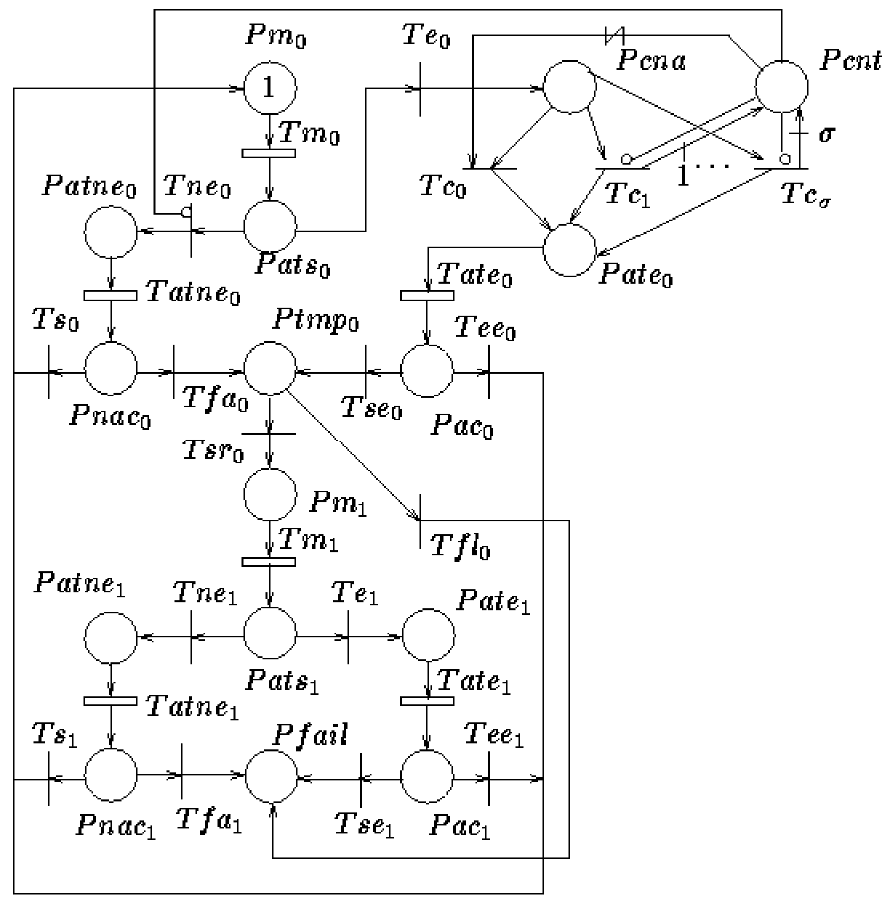


Software Fault Tolerance

- Recovery block:
  - common-mode software failures,
  - common-mode acceptance test failures,
  - failures clustered in the input stream
- A generalized model for recovery block strategy to capture the dependencies:
  - operate input within a recovery block execution
  - evaluate module output in the recovery block
  - difficult inputs clustered in input space

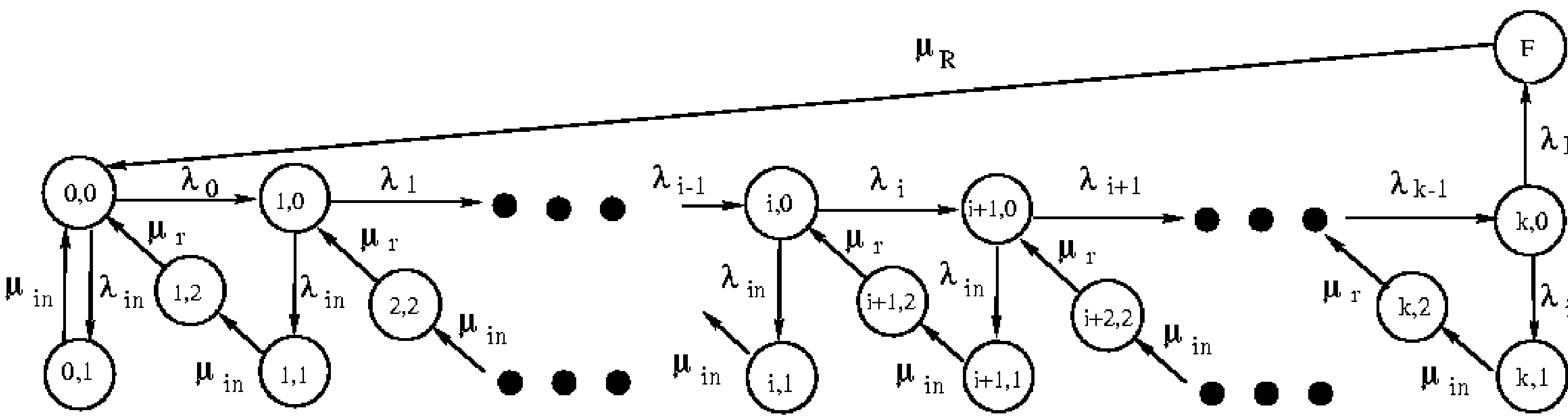
SRN Model for Recovery Block

A recovery block
with clustered failures

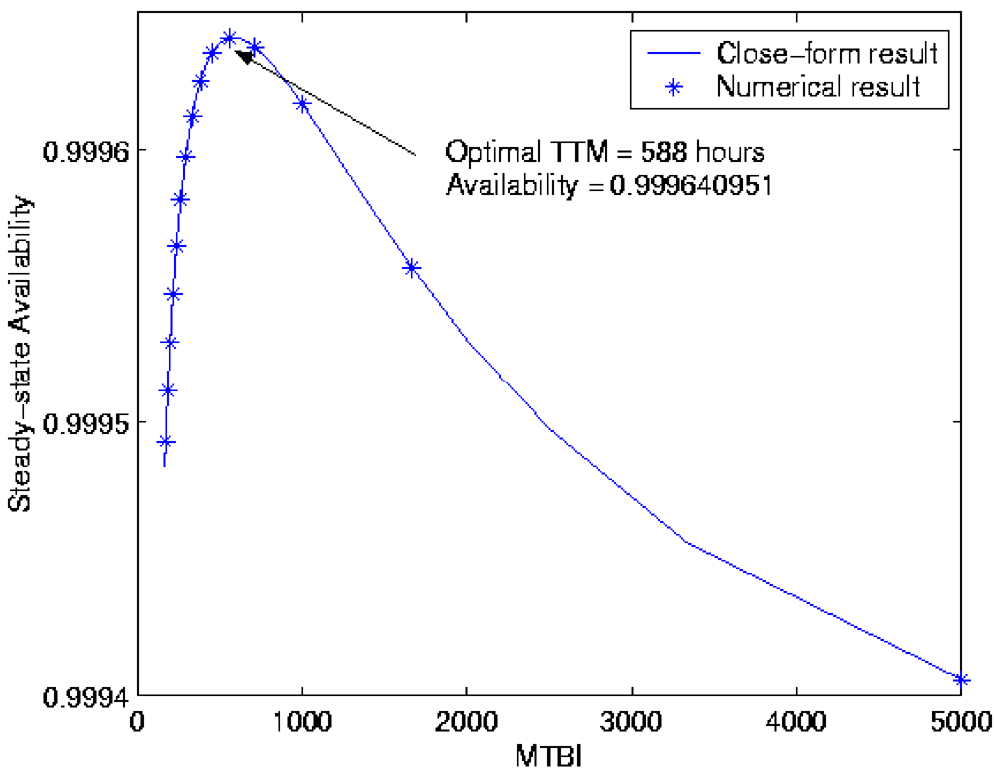


Hardware Maintenance

- Conditioned based maintenance:
  - CTMC: closed form solution
  - SHARPE: numerical solution
  - Find optimal inspection interval



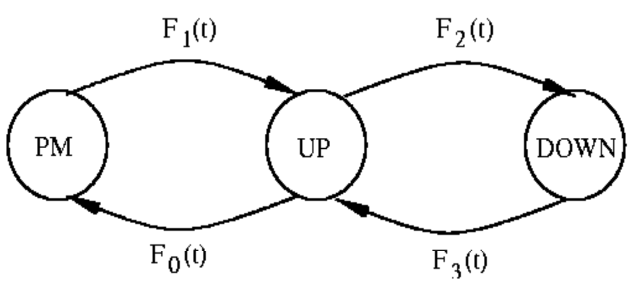
Hardware Maintenance

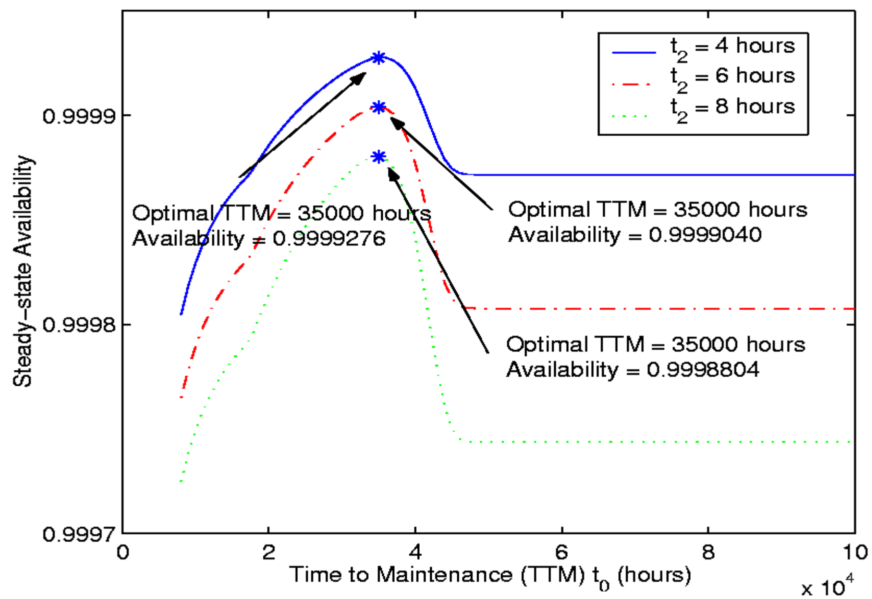


Comparison:
Closed form result
& Numerical result

Hardware Maintenance

- Time based maintenance:
  - SMP based solution (PM: preventive maintenance)
  - Find optimal maintenance interval





Availability Models

Availability models commonly used in practice assume that times to outage and recovery are exponentially distributed.

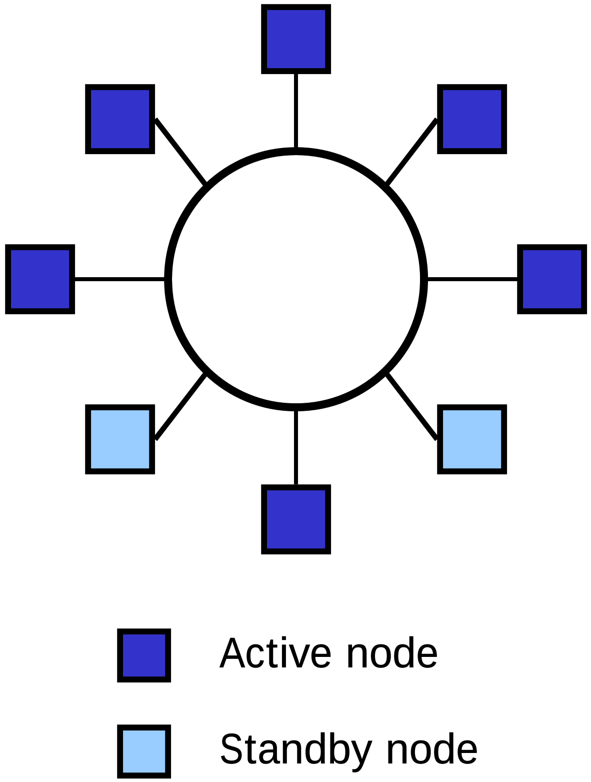
- How accurate will the all-exponential models be for systems w/ limited information of outage/recovery?
- Can we give availability bounds for such systems?

Result of a General Model

- For a system with multiple types of outage-recovery (non-exponentially distributed outrages), the underlying stochastic process is a semi-Markov process (SMP).
- We give a closed-form formula of system availability.
- Findings –
  - Only the mean value of time-to-recovery (E[TTR]) affects the availability of systems. The distribution does not matter.
  - However, the distribution of Time-to-outage (TTO) does affect availability.

Upgrade with Redundancy

- Load-sharing clustering is common in networks (wired, wireless, optical) and server systems.
- How to take advantage of the cluster structure and redundancy to upgrade hardware and software components?
- How to quantify system performance for different upgrade schemes?



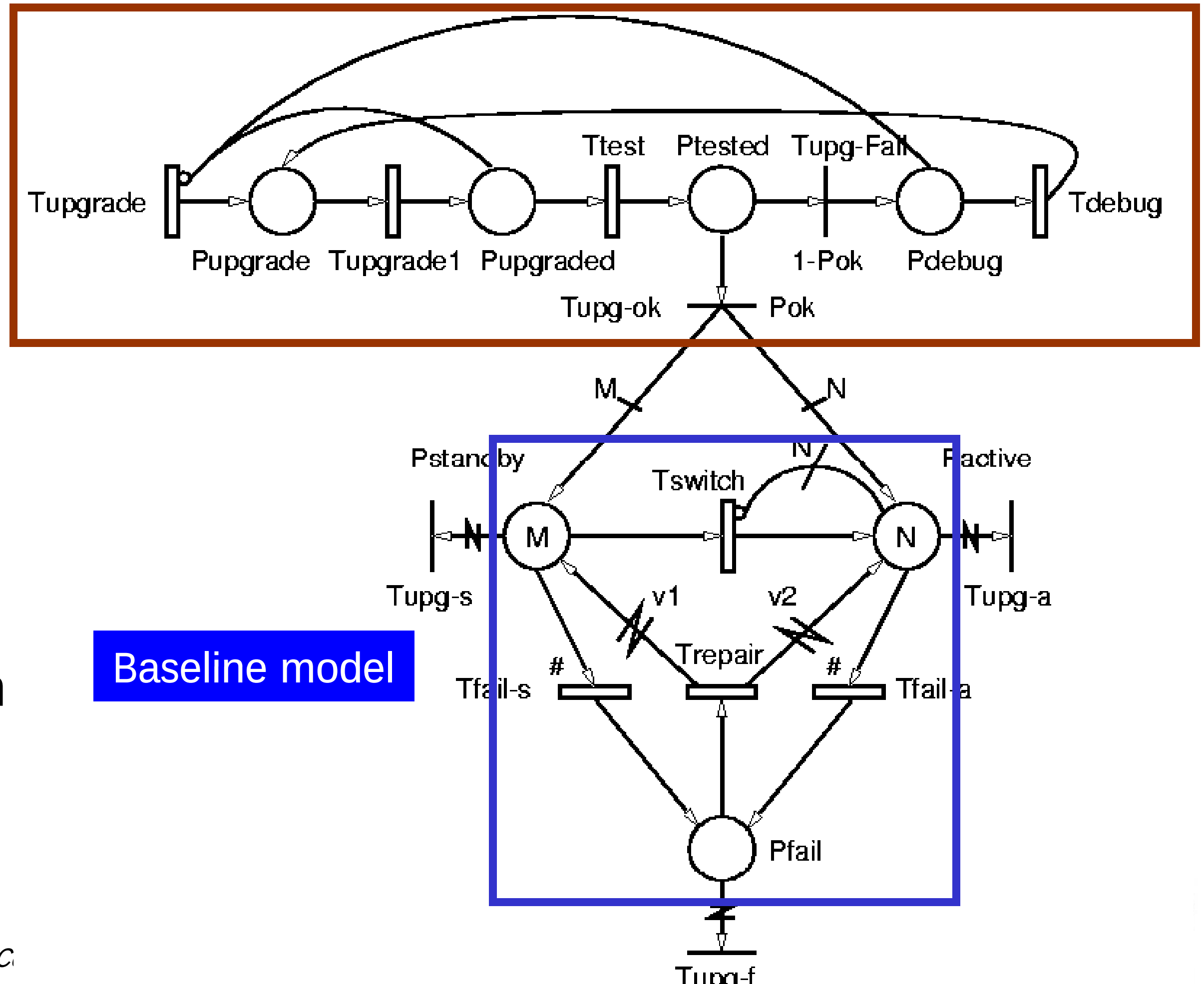
Upgrade Schemes

- <u>Direct transfer:</u>

Simple, but long downtime, for non-realtime-critical system.

- <u>Phased upgrade:</u>

Almost zero downtime, upgrade paradox, version incompatibility, etc.

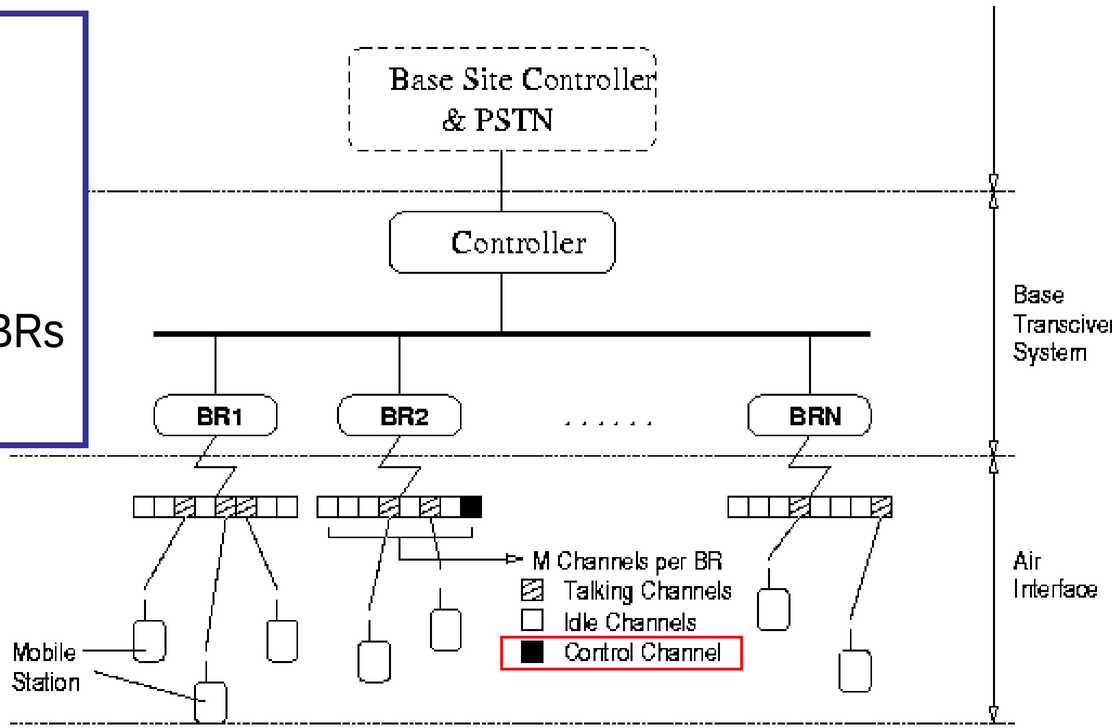


Performability of Cellular Control Channel Protection

- Each cell has N_b base repeaters (BR)
- Each BR provides M TDM channels
- One control channel resides in one of the BRs

Control channel down

System down(!)



Automatic Protection Switch

Upon *control_down*, the failed control channel is automatically switched to a channel on a working base repeater.

↓

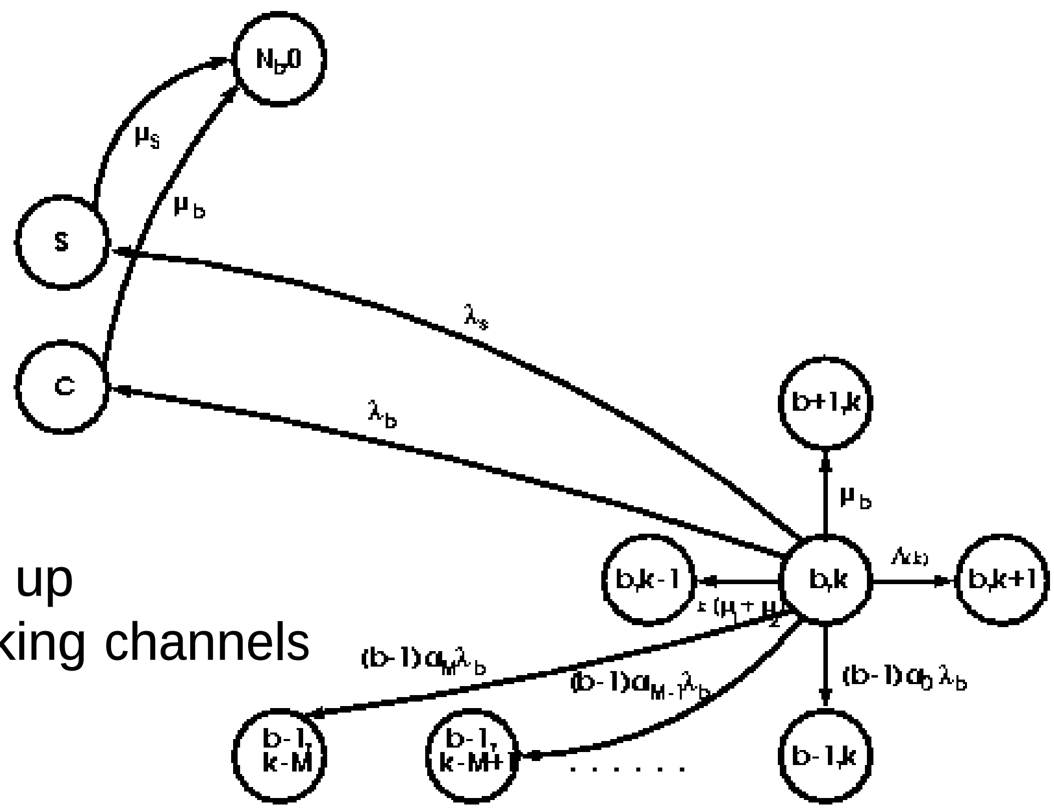
control_down causes only system *partially* down

no longer a full outage!

↓

APS A_{sys} ↑ P_{block} ↓ P_{drop} ↓

Model of System w/o APS

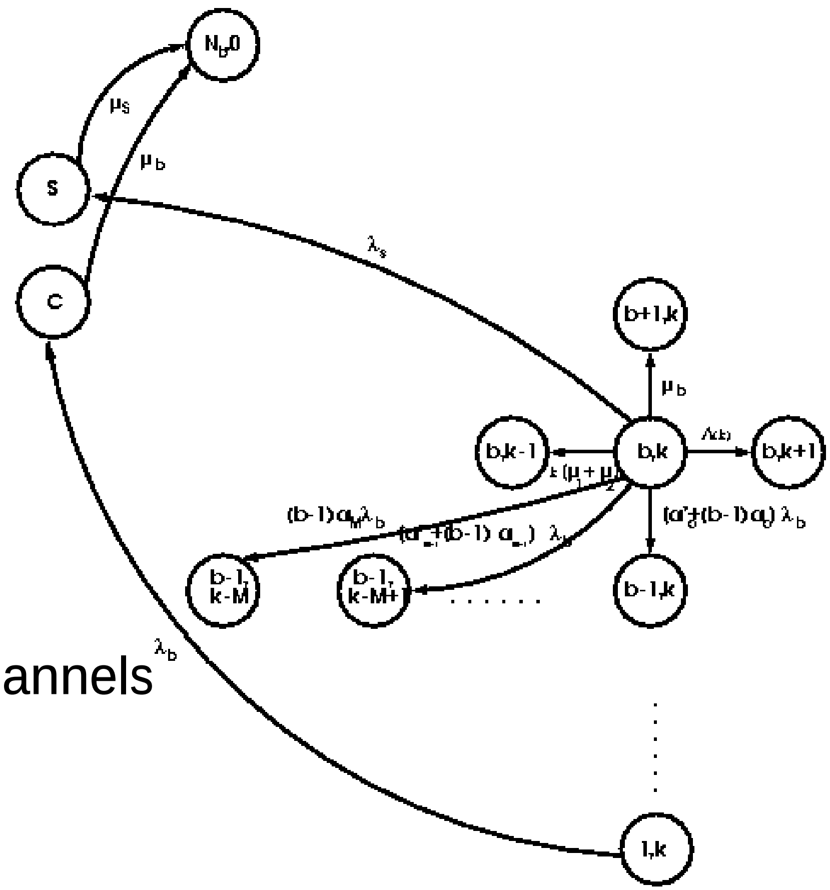


CTMC State (b,k)

b=No. of BR up

k=No. of talking channels

Model of System w/ APS



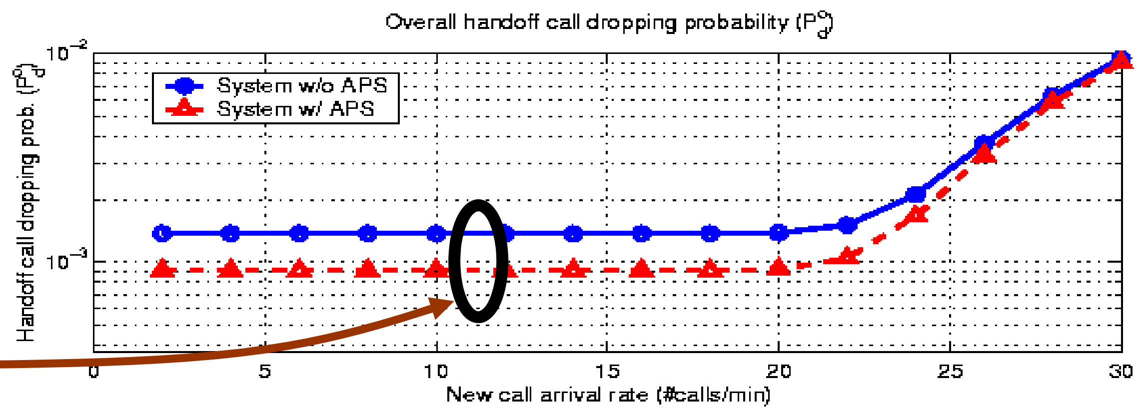
CTMC State (b,k)

b=No. of BR up

k=No. of talking channels

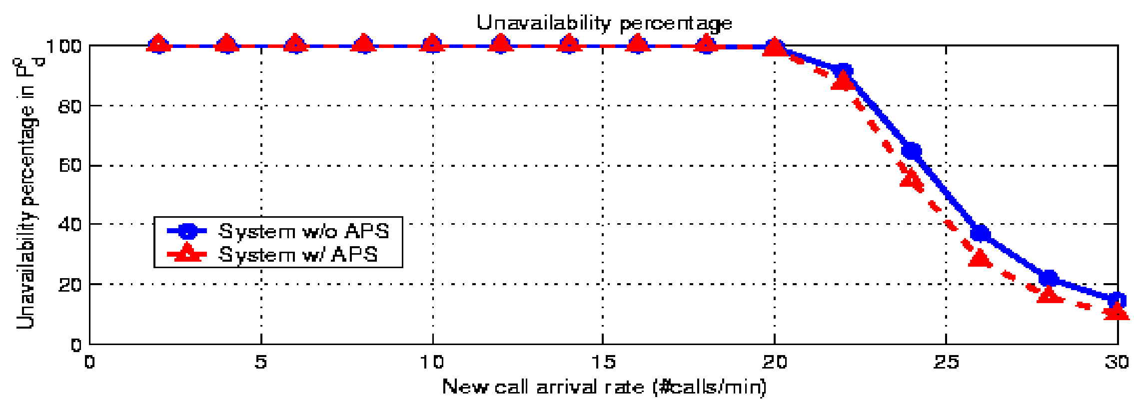
Numerical Results

Handoff Call Blocking Probability



Improvement by APS

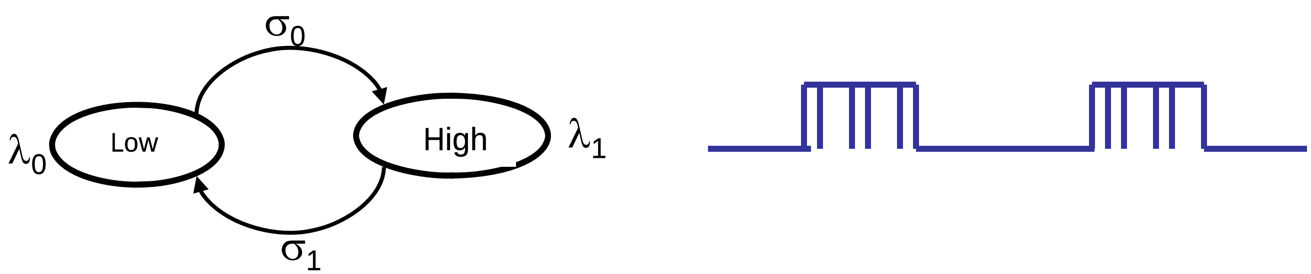
Unavailability in handoff call dropping probability



ATM Networks under Overloads

To quantify the effects of transients in ATM networks with

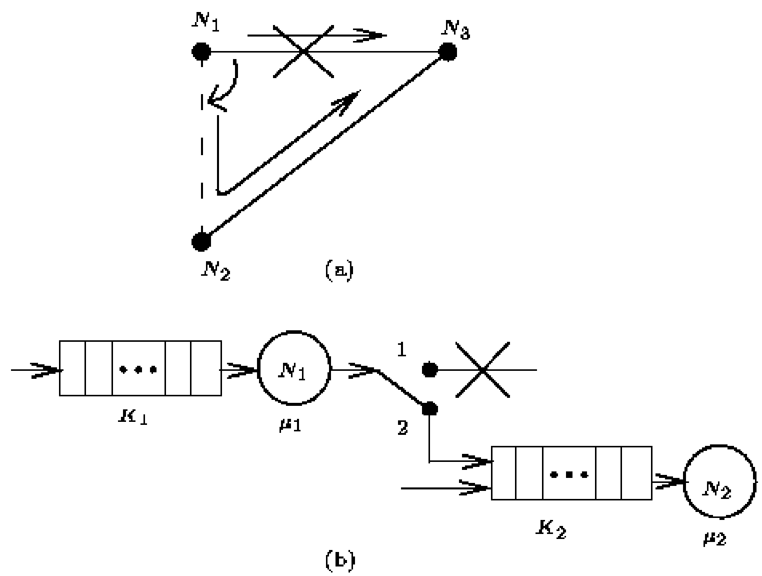
- Markov Modulated Poisson (MMPP) used to allow for Correlated arrivals
- Transient effects: relaxation time, maximum overshoot, and Expected excess number of losses in overload



2-state MMPP traffic model

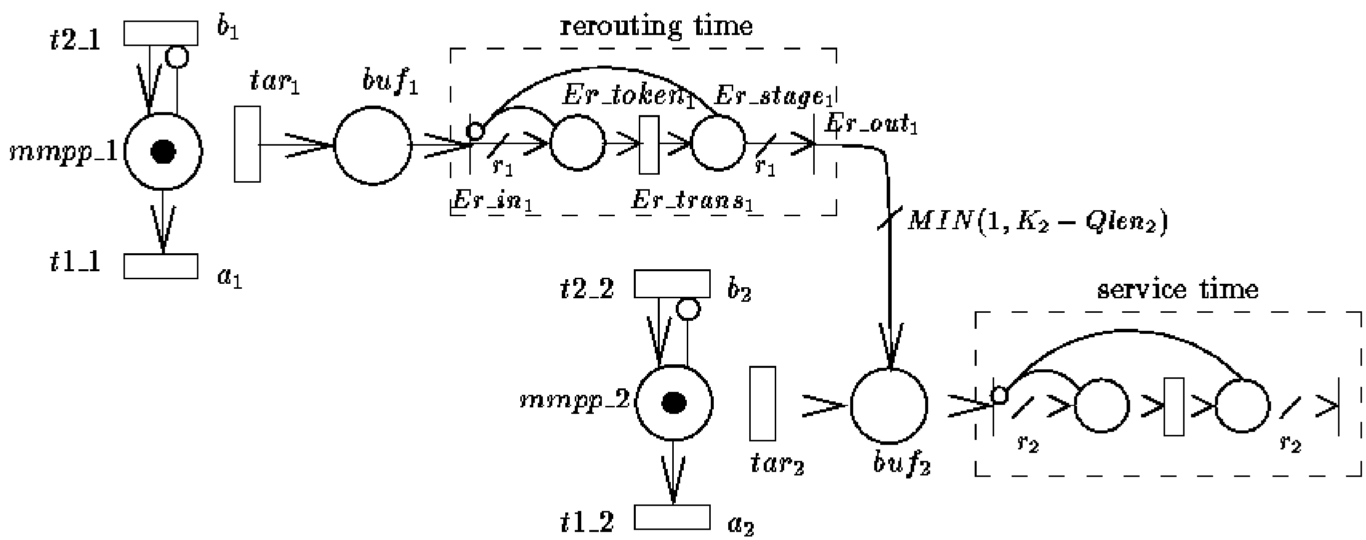
Center for Advanced Computing and Communication
Department of Electrical and Computer Engineering, Duke University

Queuing Model for ATM



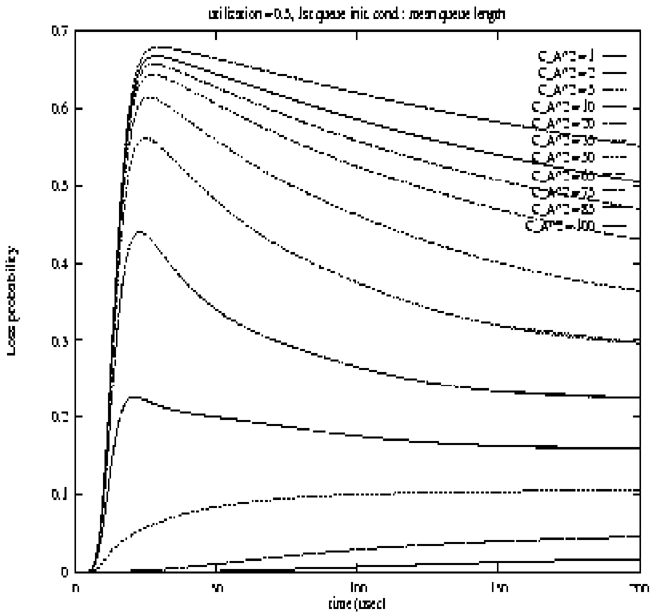
A failure occurs in the connection (N_1 and N_3).
Therefore, the traffic destined for N_3 is re-routed through N_2

SRN Model for ATM

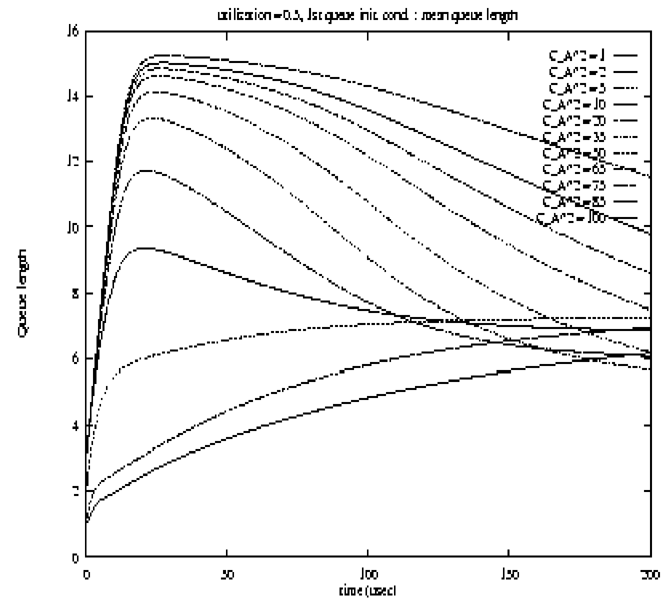


Transition	Rate Function	Enabling Function
tar_1	if (#mmpp_1) λ_1^1 else λ_2^1	$Qlen_1 < K1$
tar_2	if (#mmpp_2) λ_1^2 else λ_2^2	$Qlen_2 < K2$

Numerical Results



Loss probability vs. burstiness

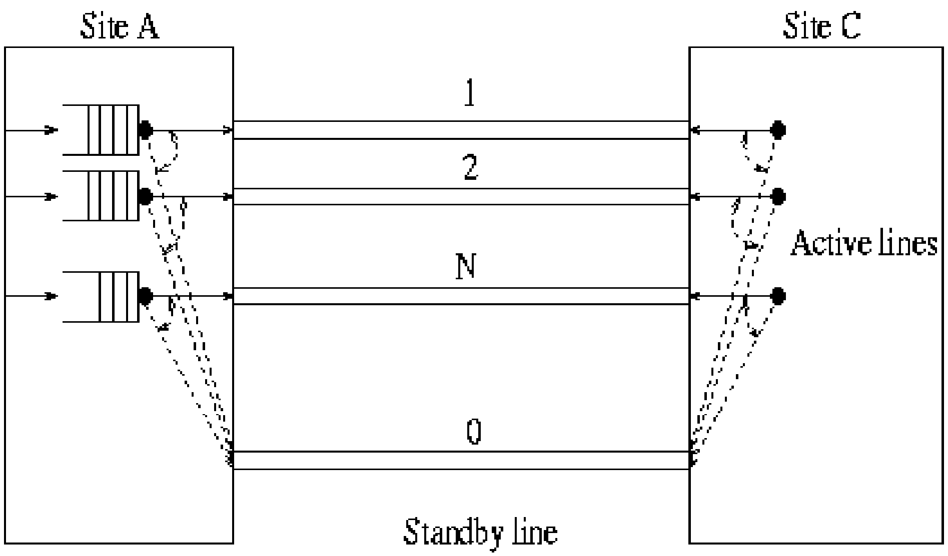


Queue length vs. burstiness

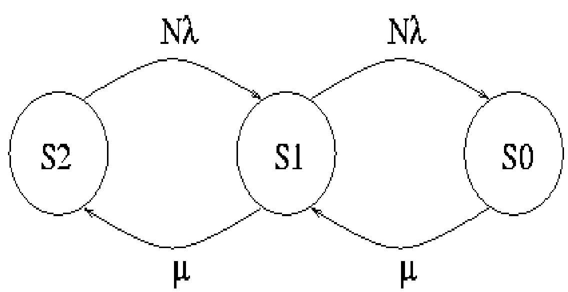
Error Recovery in Communication

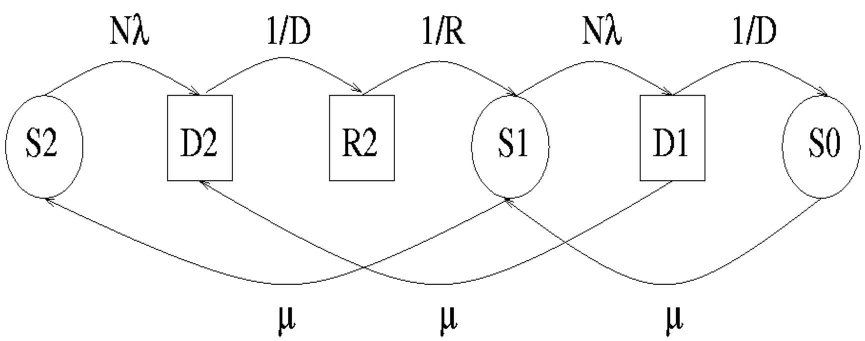
Study error recovery in communication networks using (non-exponential detection/restoration times) MRGP

1:N protection switching:



Two Modeling Methods

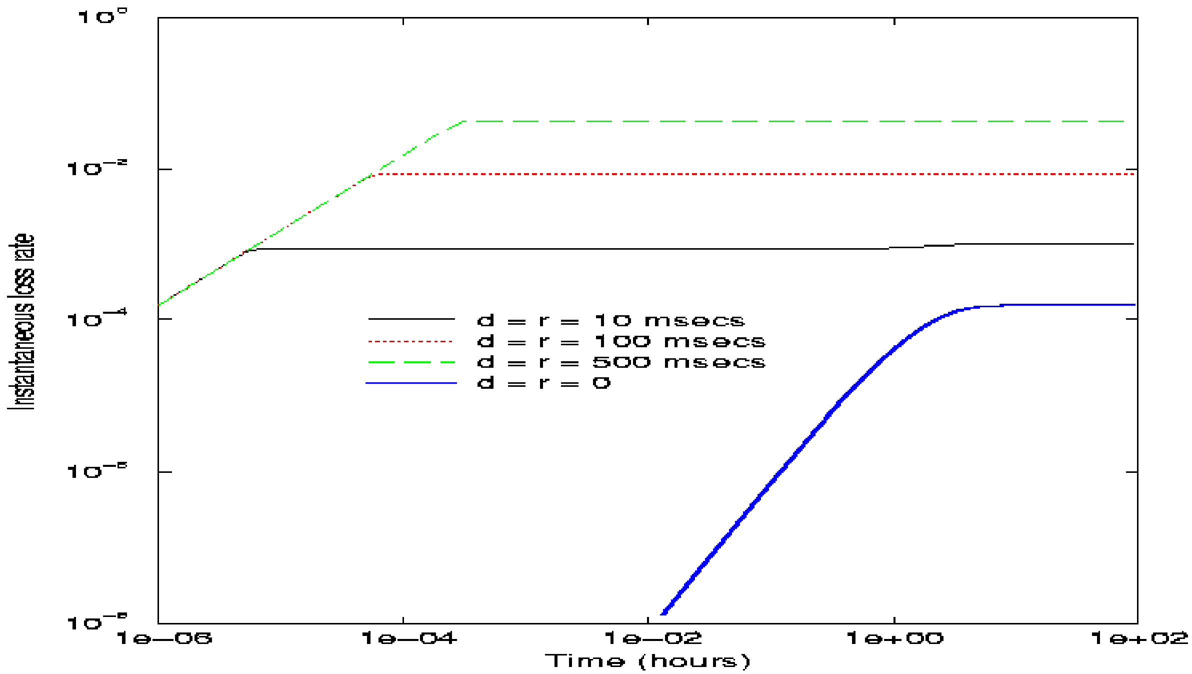




CTMC is used with ignoring the detection and restoration times

MRGP (a non-Markovian modeling) is used with including the detection and restoration times

Numerical Result



Comparison of CTMC and non-Markovian model

Open Problems

- Theoretical studies in the modeling and analysis of complex systems and failures
  - Capabilities, limitations, and relationships among formalisms such as FSPN and HS, FSPN and SDE.
- Fast algorithms
  - Fast algorithms for FSPN
  - Fast algorithms for discrete-event simulation
  - Fast algorithms for non-Markovian models
- Applications
  - Software: further exploration of software rejuvenation
  - Performability analysis of (wired and wireless) networks

The End

Thank you!